A THREE-LAYERED FRAMEWORK: AN AI GOVERNANCE GUIDE FOR GLOBAL POLICYMAKERS

Jiawei Zhang[*] *Boris Paal*[**]

Abstract

The article designs a three-layered framework to address the pressing need for more systematic and scientific methodologies in global AI governance. This framework sequentially employs Market-Invigorating Strategies, Value-Directed Rules, and Procedural Controls to achieve governance that is agile, robust, and proportionate. The first layer prioritizes market mechanisms by addressing market failures and fostering competition through consumer empowerment and competition adjustment tools. The second layer establishes value-directed guardrails through command-and-control rules as well as rights-and-responsibilities legislation when market forces prove insufficient to avoid race-to-the-bottom that may harm fundamental human values. The third layer implements procedural controls through internal management requirements and information duties to ensure compliance with the first two layers.

Throughout this process, regular risk assessment and regulatory adjustments are essential to maintain the framework's responsiveness to rapidly evolving AI technologies. This layered approach helps to achieve three core governance objectives: agility in responding to technological change, robustness in protecting fundamental values, and

[*] Ph.D. Researcher at the Technical University of Munich; Guest Researcher at Max Planck Institute for Innovation and Competition. M.Phil. (Oxon); LL.M. (U.C. Berkeley). [*Publication records here*]

[**] Full Professor of Law and Regulation of Digital Transformation at the Technical University of Munich. [*Publication records here*]

proportionality in balancing innovation with risk mitigation. Ultimately, this methodological framework enables global policymakers to craft their own customized AI governance frameworks by scientifically evaluating and harmonizing various regulatory tools while maintaining flexibility to respond to rapidly evolving AI technologies and associated risks. It will also facilitate global regulatory experimentation and international AI legislative process.

Keywords: *AI Governance; AI Regulatory Tools; AI risks; Market Mechanism; Value-Directed Rules; Procedural Controls*

Table of Contents

I. INTRODUCTION....3
II. FIRST LAYER: MARKET-INVIGORATING STRATEGIES....9
A. Consumer Empowerment....15
(a) Disclosure-to-Consumer....16
(b) AI Literacy....17
(c) Right-and-Responsibility....19
B. Competition Adjustment....21
(a) Open Source....21
(b) Interoperability....23
(c) Data Access....25
(d) Free Labor Flow....27
(e) Regulatory Sandbox....28
III. SECOND LAYER: VALUE-DIRECTED RULES....30
A. Command-and-Control....33
(a) Standard Setting....35
(b) Licensure....38
(c) Regulatory Sandbox....40
B. Right-and-Responsibility....42
IV. THIRD LAYER: PROCEDURAL CONTROLS....43
A. Internal Management....44

B. Information Duty 46
(a) Documentation 46
(b) Disclosure-to-Regulator 47
V. RISK ASSESSMENT AND REGULATION ADJUSTMENTS 50
VI. TOWARDS AGILE, ROBUST, AND PROPORTIONATE AI GOVERNANCE 53
VII. CONCLUDING REMARKS 57
APPENDIX 59

I. INTRODUCTION

U.S. Vice President J.D. Vance's speech at the Artificial Intelligence Action Summit in Paris has underscored the divergent paths taken by the United States and the European Union in regulating artificial intelligence.[1] This divergence highlights the pressing global need for designing a regulatory framework and methodology that can address broad concerns across different jurisdictions, which is the main task of this Article. Indeed, policymakers and researchers have already recognized various risks posed by AI technologies.[2] In response, worldwide jurisdictions are grappling with the complexities of AI governance within their existing legal structures.[3] Going with the tide, policymakers and researchers have designed various regulatory tools to

[1] *See* J.D. Vance, *Remarks by the Vice President at the Artificial Intelligence Action Summit in Paris, France*, AM. PRESIDENCY PROJECT (Feb. 11, 2025), https://www.presidency.ucsb.edu/documents/remarks-the-vice-president-the-artificial-intelligence-action-summit-paris-france.

[2] *See, e.g.*, Christiane Wendehorst, *Liability for Artificial Intelligence: The Need to Address Both Safety Risks and Fundamental Rights Risks*, *in* THE CAMBRIDGE HANDBOOK OF RESPONSIBLE ARTIFICIAL INTELLIGENCE: INTERDISCIPLINARY PERSPECTIVES 187, 187–92 (Oliver Mueller et al. eds., 2022) (classifying AI risks into physical and social dimensions).

[3] *See, e.g.*, Urs Gasser, *Governing AI with Intelligence*, 40 ISSUES SCI. & TECH 36, 36–39 (2024) (examining national approaches to AI governance).

mitigate these risks.[4] The European Union (EU) has pioneered its comprehensive AI legislation with the AI Act.[5]

However, policymakers tend to take a *nirvana* approach, intuitively assuming that their proposals can lead the world from an imperfect reality to an idealized improved state while committing three old fallacies: "the grass is always greener on the other side," "free lunch," and "people could be different" fallacies.[6] This overestimation of the legislators' ability and their—even sometimes unrealistic—proposals is largely derived from insufficient research in designing a comprehensive regulatory framework and the evaluation benchmarks at the macro level to compare the effectiveness of different regulatory proposals, taxonomize these tools by their functions, and orchestrate their synergistic deployment. Without clear guidance of systematic regulatory methodologies, researchers and regulators are prone to propose imbalanced, under-evaluated, and short-sighted rules. This not only fails

---

[4] *See, e.g.*, Urs Gasser, *Navigating AI Governance as a Normative Field: Norms, Patterns, and Dynamics*, *in* REALIZING THE PROMISE AND MINIMIZING THE PERILS OF AI FOR SCIENCE AND THE SCIENTIFIC COMMUNITY 57–92 (Kathleen H. Jamieson et al., ed., 2025) (establishing a systematic framework for various AI regulatory instruments and mechanisms); *see also* Orly Lobel, *The AI Regulatory Pyramid: A Taxonomy & Analysis of the Emerging Toolbox in the Global Race for the Regulation and Governance of Artificial Intelligence* 19–39 (San Diego Legal Studies Paper No. 24-008, 2024); Margot E. Kaminski, *Regulating the Risks of AI*, 103 BOS. U. L. REV. 1347, 1369–89 (2023); Martin Ebers, *Truly Risk-Based Regulation of Artificial Intelligence How to Implement the EU's AI Act*, EUR. J. RISK REGUL. 1 (2024).

[5] *See* Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828, 2024 O.J. (L) 1 (hereinafter EU AI Act or AI Act).

[6] Demsetz raised the *nirvana approach* of policymaking and its fallacies. *See* Harold Demsetz, *Information and Efficiency: Another Viewpoint*, 12 J.L. & ECON. 1, 2–7 (1969). For nirvana AI governance and three fallacies in AI policymaking, see Jiawei Zhang, *Nirvana AI Governance: How AI Policymaking is Committing Three Old Fallacies*, 15 THE REGULATORY REVIEW IN DEPTH (forthcoming 2025).

to satisfactorily resolve the pressing issues but may even cause more side effects in the long term.

At a macro level, various regulatory actors simultaneously work on a specific regulatee. In Lawrence Lessig's words, regulatory actors include market, law, architecture, and norm.[7] They collectively and interactively act on the "pathetic dot" (regulatee).[8] Lessig used various examples to illustrate how law interacts with the other three modalities to regulate the regulatees. Taking the seatbelt as an example, to mitigate the risks of injuries and fatalities from traffic accidents, law can directly regulate behavior by requiring drivers (the "dot") to wear seatbelts; law can regulate social norms to indirectly regulate people's behaviors by funding education campaigns to foster societal disapproval towards those who do not wear seatbelts; law can regulate the market to incentivize the drivers to wear the seatbelt by subsidizing insurance companies to offer reduced rates to seatbelt wearers; and, law can also regulate the architecture by mandating automatic seatbelts of the automobiles.[9]

Lessig's pioneering formula is also applicable to AI governance. For example, law can regulate the use of AI products by setting redlines and penalties (*i.e.*, "dot" regulation);[10] law can regulate social norms by supporting AI literacy training programs (*i.e.*, norm regulation);[11] law can regulate the market by incentivizing the AI market players to compete for

[7] Lessig calls them regulation modalities. *See* LAWRENCE LESSIG, CODE 122–23 (2006).

[8] *Id.*

[9] *Id.*

[10] *See generally* Miles Brundage et al., *The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation* (Feb. 2018), https://arxiv.org/pdf/1802.07228 (focusing on regulating malicious use of AI to enhance the security of AI).

[11] *See, e.g.*, Teresa Scantamburlo et al., *Artificial Intelligence across Europe: A Study on Awareness, Attitude and Trust*, IEEE TRANSACTIONS ON A.I. 1, 1 (2024) (arguing that "[t]he introduction of rules that ensure legal and ethical standards, along with the activity of high-level educational entities, and the promotion of AI literacy are identified as key factors in supporting a trustworthy AI ecosystem").

delivering ethical products (*i.e.*, market regulation); [12] law can also regulate architecture by mandating AI companies to embed ethical values into the development of AI technologies (*i.e.*, code/architecture regulation). [13] Regulators and researchers have designed a myriad of governing measures that can be classified into the foregoing four modalities.[14] While these four regulatory actors provide a comprehensive classification system for AI governance, it is still crucial to further design a scientific policymaking guideline to categorize and harmonize these four modalities to respond to the rapidly evolving nature of AI and other cutting-edge technologies and the derived risks, in which way the entire governance framework tackles and addresses agility, robustness, and proportionality.[15]

This research is a methodological study of AI regulation, establishing a normative guideline for AI risk governance at the macro level and providing detailed prescriptions of each regulatory toolbox and regulatory tool at the micro level. The ultimate objective of this research is to generate an AI governance framework that is agile, robust, and proportionate to handle unpredictable risks and uncertainties brought about by emerging technologies. Thus, this article suggests a three-layered framework to fulfill this goal. Following this framework, policymakers can sequentially select various regulatory instruments from

[12] *See generally* Jiawei Zhang, *Regulating Chatbot Output via Inter-Informational Competition*, 22 NW. J. TECH. & INTELL. PROP. 109 (2024) (emphasizing the effectiveness of the market mechanism in AI-enabled chatbot output regulation).

[13] *See generally* Andreas Tsamados et al., *The Ethics of Algorithms: Key Problems and Solutions*, 37 AI & SOC'Y 215 (2022); Iason Gabriel, *Artificial Intelligence, Values, and Alignment*, 30 MINDS & MACHS. 411 (2020).

[14] For an excellent classification of current AI governance models, see Gasser, *supra* note 4.

[15] For relevant research in this direction, see, for example, Julia Black & Andrew Murray, *Regulating AI and Machine Learning: Setting the Regulatory Agenda*, 10 EUR. J.L. & TECH. 21 (2019).

three-layered toolboxes, namely *Market-Invigorating Strategies*, *Value-Directed Rules*, and *Procedural Controls*.

> ***1st Layer – Market-Invigorating Strategies****: designing governance strategies to invigorate market competition by fixing market failures, such as empowering AI consumers and preventing market concentration and anti-competitive conduct, to incentivize AI market players to voluntarily improve their services and de-risk their systems.*
>
> ***2nd Layer – Value-Directed Rules:*** *imposing mandatory bottom-line rules on AI operators and consumers, requiring them to develop and use AI products/services in alignment with specific human values (including fundamental rights and freedoms), to directly address concerns that the AI market might fail to address.*
>
> ***3rd Layer – Procedural Controls****: requiring self-management and information disclosure, and auditing AI operators and conducting iterative risk assessments to ensure that their operational procedures are complying with the 1st and 2nd layers' rules and that risks are effectively controlled to an acceptable extent.*

To accomplish this task, this article proceeds in three main parts. Parts II to IV provide key understandings of the proposed three-layered AI governance framework, delineating and justifying the three-layered arrangements. In each part, we equip each layer with various regulatory tools, each accompanied by a detailed prescription. Part V highlights the risk assessment throughout the entire process of policymaking and generates some guidance on regulation adjustments. Part VI explains the three core values in AI governance, namely agility, robustness, and proportionality, that the proposed three-layered framework embodies and promotes. Part VII concludes by reiterating the practical and

theoretical contributions that this research will make. At the end of this article, we generate a user-friendly checklist as an AI governance guide for policymakers to understand and implement.

To facilitate comprehensibility, certain concepts we used throughout the article should be clarified upfront. In this article, some key terminologies are used in alignment with the definitions set by Article 3 of the EU AI Act.[16] Specifically, we employ "risk" to indicate "the combination of the probability of an occurrence of harm and the severity of that harm,"[17] and "AI operator" to denote all parties involved in an AI system, including AI providers, deployers, importers, and distributors.[18] Additionally, "governance" (and "govern[ing]") in this article indicates a governmental policymaking process that aims at effectively addressing a series of problems. And, we use "regulation" (and "regulat[e/ing]") in a broad way to comprise all tools of risk mitigation, including market, norm, and law, despite regulation sometimes being narrowly used as the antithesis of market competition. Accordingly, the concept "regulatee" includes both AI operators and consumers of AI products (or services).

Please note that this article serves merely as a general guideline—presenting the pros and cons of regulatory proposals in a neutral and balanced way—and does not discount any specific regulatory tool, since the effectiveness of a tool is contingent on many factors, including but not limited to market conditions, consumers' literacy, level of risk tolerance, goals to be achieved, and jurisdiction-specific cultural background.[19] Also, our prescriptions will not be exhaustive. Especially for those commonly used and well-known regulatory tools, we will provide readers with some

[16] *See* EU AI Act, *supra* note 5, at art. 3

[17] *Id.* at art. 3(2).

[18] *Id.* at art. 3(8).

[19] *See generally* Matthew R. O'Shaughnessy et al., *What Governs Attitudes toward Artificial Intelligence Adoption and Governance?*, 50 SCI. & PUB. POL'Y 161 (2023).

useful publications that signal the latest policy focus and model policy design.

II. FIRST LAYER: MARKET-INVIGORATING STRATEGIES

Different regulatory modalities can concurrently target at a single risk, but that does not mean that policymakers may employ them randomly and arbitrarily. First and foremost, regulators should consider using the market mechanism to mitigate the recognized risks. A long-overlooked but well-established rule is that market failure constitutes the primary rationale for government regulation.[20] In other words, regulation should be largely conditioned on market failure––if the market can work effectively to realize the regulatory goals, then additional regulation would be unnecessary. Even those proponents of regulation in specific fields would first present market failures before proposing their regulatory proposals. [21] This means that among the four regulatory modalities, the market takes precedence over law (or regulation).[22] The

[20] *See* A. C. PIGOU, THE ECONOMICS OF WELFARE, pt.II (1920); Joseph E. Stiglitz, *Government Failure vs. Market Failure: Principles of Regulation*, *in* GOVERNMENT AND MARKETS (Edward J. Balleisen & David A. Moss eds., 2009); ROBERT BALDWIN, MARTIN CAVE & MARTIN LODGE, UNDERSTANDING REGULATION: THEORY, STRATEGY, AND PRACTICE 15–22 (2nd ed. 2012); W. KIP VISCUSI, JOSEPH E. HARRINGTON JR. & DAVID E. M. SAPPINGTON, ECONOMICS OF REGULATION AND ANTITRUST 3–4 (5th ed. 2018).

[21] *See, e.g.*, Kenneth J. Arrow, *Uncertainty and the Welfare Economics of Medical Car*e, 53 AM. ECON. REV. 345 (1978); WILLIAM J. BAUMOL & WALLACE E. OATES, THE THEORY OF ENVIRONMENTAL POLICY (2nd ed. 2005); Adam B. Jaffe, Richard G. Newell & Robert N. Stavins, *A Tale of Two Market Failures: Technology and Environmental Policy*, 54 ECOLOGICAL ECON. 164 (2005); Tim Wu, *Network Neutrality, Broadband Discrimination*, 2 J. ON TELECOMM. & HIGH TECH. L. 141 (2003); Lina M. Khan, *Sources of Tech Platform Power*, 2 GEO. L. TECH. REV. 325 (2018).

[22] Here, to clarify, we are not discussing the relationship between law and the market, and we are not saying that the market precedes law in the temporal sense. We fully agree that law makes the market possible—"it is law that enforces contracts, establishes property, and regulates currency." *See* LAWRENCE LESSIG, CODE 127 (Version 2.0, 2006). It is with this awareness that we prioritize the market invigoration by laws. We are

underlying reason for prioritizing market mechanism lies in a long-recognized belief that the market acting as an invisible hand is better at utilizing decentralized information in allocating social resources relative to government intervention.[23] This lesson is particularly correct in mitigating risks arising from rapidly changing technologies and evolving realities.

Therefore, before proposing a sequence of ambitious regulatory tools, policymakers should first evaluate the market conditions to see whether the market can and does work effectively in mitigating the recognized risks. For instance, in the 1960s, the U.S. Federal Communications Commission (FCC) of the United States adopted the mandatory fairness doctrine to require radio broadcast stations to present controversial viewpoints in a fair and balanced way, to ensure that divergent and competing voices can be heard by the public, against the background of the physical scarcity of broadcast frequency.[24] However, the advent and prevalence of various competing information technologies, like cable and satellite, made the scarcity rationale invalid and the fairness doctrine unnecessary.[25] Similarly, 40 years later, the FCC used network

---

arguing for a thinking pattern that can be used in practice—when policymakers design new policies to govern emerging risks, they should first consider market tools and evaluate the market performance, and then tailor their ethics-directed rules.

[23] *See* ADAM SMITH, THE WEALTH OF NATIONS bk. IV, Ch. II & XI (Liberty Press 1979) (1776); F. A. Hayek, *The Use of Knowledge in Society*, 35 AM. ECON. REV. 519 (1945); MILTON FRIEDMAN & ROSE FRIEDMAN, FREE TO CHOOSE: A PERSONAL STATEMENT (1980).

[24] Applicability of the Fairness Doctrine in Handling of Controversial Issues of Public Importance, 29 Fed. Reg. 10415 (1964) (proposing fairness doctrine to require broadcast licensees to present these controversial public issues in a fair and balanced way so as to ensure a free and fair competition of competing views); Red Lion Broad. Co. v. F.C.C., 395 U.S. 367, 390 (1969) ("Because of the scarcity of radio frequencies, the Government is permitted to put restraints on licensees in favor of others whose views should be expressed on this unique medium.").

[25] Loveday v. F.C.C., 707 F.2d 1443, 1459 (D.C. Cir. 1983) (explaining that the First Amendment protections of broadcast political speech may expand due to the dramatic increase in the number of broadcast stations); F.C.C. v. League of Women Voters of California, 468 U.S. 364, 376 (1984) ("[W]ith the advent of cable and satellite television

neutrality rules—namely no-blocking, no-throttling, no-paid-prioritization, and transparency rules—to ensure a fair and pluralistic internet market.[26] However, in 2018, the FCC eventually revocated network neutrality rules and only maintained the transparency rules after finding that the broadband internet industry can self-regulate under sufficient market forces.[27]

To that end, policymakers should identify where market failures are and find solutions to fix these failures accordingly.[28] The common causes of market failures include but are not limited to natural monopoly, concentrated market, scarcity, externalities, information asymmetries, and consumer irrationality.[29] When the market mechanism fails, the top priority for policymakers is to design some regulatory strategies to fix these market failures, provide proper incentives, invigorate the market, and enable competition, instead of imposing harsher value-directed rules and controls. In addition to competition law (or antitrust law) as a form of bottom-line and ex-post rules, some ex-ante interventions can also boost market competition. For example, when there is a natural monopoly, the appropriate remedy is to impose excess-profits tax, rather than to set entry restrictions and price caps; [30] when the market is highly concentrated, one of the strategies is to lower the market barriers by providing some economic support for start-ups or setting mandatory

technology, communities now have access to such a wide variety of stations that the scarcity doctrine is obsolete.").

[26] In the Matter of Preserving the Open Internet Broadband Indus. Pracs., 25 F.C.C. Rcd. 17905 (2010).

[27] In the Matter of Restoring Internet Freedom, 33 F.C.C. Rcd. 311 (2018).

[28] *See* Mark A. Lemley, *Free the Market: How We Can Save Capitalism from the Capitalists*, (manuscript at 3) (Jan. 11, 2024), https://papers.ssrn.com/abstract=4692267 ("The market needs to be tempered with regulation aimed at its imperfections.").

[29] *See, e.g.*, BALDWIN, CAVE & LODGE, *supra* note 20, at 15–22.

[30] *See, e.g.*, RICHARD A. POSNER, NATURAL MONOPOLY AND ITS REGULATION 110 (1999).

obligations of sharing essential facilities;[31] when there are externalities, the spillover costs can be internalized by taxing the company which causes the externalities[32] or by adjusting education and offering information in order to nudge consumers to consider these spillover effects when making purchasing decisions; [33] when information asymmetry exists between marketers and consumers, mandatory disclosure may be appropriate to ensure consumers are well-informed before their decision-making;[34] just to name a few.

Some market failures exist in the AI industries. First, information asymmetry between AI giants and consumers may hinder consumers from making rational decisions and undermine the market mechanism. The workings of AI are inherently opaque to those who are not technical experts while AI-made decisions, especially those in healthcare, finance,

---

[31] *See, e.g.*, Lemley, *supra* note 28.

[32] *See* PIGOU, *supra* note 20.

[33] This method acts on the market by reinforcing the third modality—norms. For examples, see Karine Nyborg, Richard B. Howarth, & Kjell Arne Brekke, *Green Consumers and Public Policy: On Socially Contingent Moral Motivation*, 28 RES. & ENERGY ECON. 351 (2006); Fredrik Carlsson et al., *The Use of Green Nudges as an Environmental Policy Instrument*, 15 REV. ENV'T ECON. & POL'Y 216 (2021); Christian Schubert, *Green Nudges: Do They Work? Are They Ethical?* 132 ECOLOGICAL ECON. 329 (2017); Eric Cardella, Bradley T. Ewing & Ryan B. Williams, *Green Is Good—The Impact of Information Nudges on the Selection of Voluntary Green-Power Plans*, 43 ENERGY J. 1 (2022).

[34] *See generally* Cary Coglianese, Richard J. Zeckhauser & Edward Parson, *Seeking Truth for Power: Informational Strategy and Regulatory Policy Making*, 89 MINN. L. REV. 277 (2004); Bradley C. Karkkainen, *Information as Environmental Regulation: TRI and Performance Benchmarking, Precursor to a New Paradigm?*, 89 GEO. L.J. 257 (2001); Paul R. Kleindorfer & Eric W. Orts, *Informational Regulation of Environmental Risks,* 18 RISK ANALYSIS 155 (1998); Mary L. Lyndon, *Information Economics and Chemical Toxicity: Designing Laws to Produce and Use Data*, 87 MICH. L. REV. 1795 (1989); *see also* Cass R. Sunstein, *Informational Regulation and Informational Standing: Akins and Beyond*, 147 U. PA. L. REV. 613, 673 (1999) ("[I]nformational strategies attempt to improve the ordinary operations of markets by ensuring that workers and consumers are adequately informed.").

education, and judicial rulings, have profound implications.[35] Consumers have limited knowledge of when, what, and how AI models are used, which leads to difficulty in comparing different AI-driven products or services and selecting a better one.[36] As the "market for lemons" showed, when product quality is invisible to consumers, consumers' decision-making relies more heavily on other visible but monotonous factors, like price.[37] If knowledge is power, then consumers have limited power to check those powerful AI giants.[38] As a result, the market mechanism is fundamentally distorted and destined to fail.

Second, AI markets exhibit a high tendency towards market concentration. Researchers have found that various factors act as barriers to entry into the AI market. First are the financial barriers—AI companies face significant fixed costs to access infrastructure, energy, sufficient high-quality data, computing power, talents, and other sources that are essential for developing an AI model.[39] The economics of scale and network effects (e.g., data feedback loop) derived from these resources

---

[35] *See* Jenna Burrell, *How the Machine "Thinks": Understanding Opacity in Machine Learning Algorithms* 3 BIG DATA & SOC'Y 1, 4 (2016) ("Opacity as technical illiteracy.").

[36] *See generally* FRANK PASQUALE, THE BLACK BOX SOCIETY: THE SECRET ALGORITHMS THAT CONTROL MONEY AND INFORMATION (2016).

[37] *See* George A. Akerlof, *The Market for "Lemons": Quality Uncertainty and the Market Mechanism*, 84 Q. J. ECON. 488 (1970).

[38] ELI PARISER, THE FILTER BUBBLE: WHAT THE INTERNET IS HIDING FROM YOU 81 (2011) ("If knowledge is power, then asymmetries in knowledge are asymmetries in power.").

[39] *See* Jai Vipra & Anton Korinek, *Market Concentration Implications of Foundation Models: The Invisible Hand of ChatGPT*, BROOKINGS 9–19 (Center on Regulation and Markets Working Paper No. 9, Sep. 2023), https://www.brookings.edu/articles/market-concentration-implications-of-foundation-models-the-invisible-hand-of-chatgpt/; Tejas N. Narechania & Ganesh Sitaraman, *An Antimonopoly Approach to Governing Artificial Intelligence*, YALE L. & POL'Y REV. (manuscript at 8–23) (forthcoming 2024), https://papers.ssrn.com/abstract=4597080.

greatly favor the incumbents.[40] Moreover, incumbent AI companies can erase potential competition through vertical integration, strategic partnerships, and financial investments, which further entrench their positions and increase market concentration.[41] To make matters worse, some jurisdictions are racing to propose harsher standards for AI technologies, raising the market barriers and scaring off start-ups and small companies that are unable to afford the exorbitant compliance costs.[42]

Third, the development and maintenance of AI products produce strong spillover effects, especially the costs of environmental pollution and energy consumption, that are not adequately reflected in the pricing system. For example, researchers have estimated that the training of BLOOM, a large language model developed by Hugging Face, can lead to 50 metric tons of carbon dioxide emissions, comparable to around 60 flights between London and New York.[43] Even worse, OpenAI's GPT-3,

---

[40] *See* Maurice E. Stucke & Ariel Ezrachi, *Antitrust & AI Supply Chains*, (manuscript at 5–10) (2024).

[41] *Id.* at pt.III; Narechania & Sitaraman, *supra* note 39, at 6; U.S. Dep't of Justice, Fed. Trade Comm'n Joint, Eur. Comm'n & UK Competition & Mkts Auth, Statement on Competition in Generative AI Foundation Models and AI Products 2 (July 23, 2024), https://www.justice.gov/atr/media/1361306/dl?inline (Joint Statement).

[42] *See, e.g.*, Thibault Schrepel, *Decoding the AI Act: A Critical Guide for Competition Experts* 11–18 (ALTI Working Paper, 2023) (arguing that the EU AI Act may distort market competition because it is not technology neutral in nature, distributes the regulatory burden unevenly, and significantly raises the threshold of market entry); Neel Guha et al., *AI Regulation Has Its Own Alignment Problem: The Technical and Institutional Feasibility of Disclosure, Registration, Licensing, and Auditing*, 92 GEO. WASH. L. REV. (forthcoming 2024) (manuscript at 52–53), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4634443.

[43] *See* Alexandra Sasha Luccioni, Sylvain Viguier & Anne-Laure Ligozat, *Estimating the Carbon Footprint of BLOOM, a 176B Parameter Language Model*, (Nov. 3, 2022), https://arxiv.org/pdf/2211.02001; *see also* Melissa Heikkilä, *We're Getting a Better Idea of AI's True Carbon Footprint*, MIT TECH. REV. (Nov. 14, 2022), https://www.technologyreview.com/2022/11/14/1063192/were-getting-a-better-idea-of-ais-true-carbon-footprint/.

trained on less efficient hardware, was estimated to produce over 10 times the carbon dioxide emission of BLOOM.[44] In addition, the water footprint of AI models is also staggering. Researchers disclosed that GPT-3 training needs 5.4 million liters of water, and it will consume an additional 500 ml of water for around 10–50 responses.[45] GPT-4 and other models that have larger sizes may correspondingly need more water to cool down their systems.[46]

To address these potential market failures and invigorate the AI market competition, we have two toolboxes in hand, namely, *Consumer Empowerment* which bridges the power gap between the downstream consumers and upstream businesses, and *Competition Adjustment* which adjusts the competition relationship between different competing AI operators.

A. Consumer Empowerment

The consumer empowerment scheme is to enable consumers to compare different AI products more critically, make decisions more rationally, and switch more easily to another service provider when feel unsatisfied with the current one. In this way, upstream companies can feel the pressure from consumers and automatically align their models with the consumers' expectations to prevent user churn and maintain their competitive advantage. We argue that three tools can achieve these goals, namely disclosure-to-consumer rules, AI literacy programs, and right & responsibility legislations. This regulatory toolbox can fix some market failures by bridging the information asymmetry, easing the consumer flow, and internalizing the spillover effects.

---

[44] *See* Heikkilä, *supra* note 43.

[45] *See* Pengfei Li et al., *Making AI Less "Thirsty": Uncovering and Addressing the Secret Water Footprint of AI Models*, (manuscript at 3) (Oct. 29, 2023), https://arxiv.org/pdf/2304.03271.

[46] *Id.*

(a) Disclosure-to-Consumer

Disclosure rules, also called transparency rules, have long been recognized as an essential regulatory means.[47] Here, we use the term "disclosure-to-consumer" as distinguished from "disclosure-to-regulator" because they have different functions. The latter, discussed in Part IV, is more of a procedural control, through which regulators can gain the knowledge of whether or not the company is complying with the existing rules. The former, however, functions to empower consumers intellectually by offering information about the AI product or service that is necessary for consumers' decisions. For example, Sara Gerke draws lessons from "nutrition facts labels" of food industries to design AI labeling standards, arguing that "because adequate labeling can inform users, among other things, about the benefits, potential risks, and limitations of the AI/ML-based medical device, it can promote autonomous decisions on whether to use the device."[48] As Frank Pasquale puts it similarly, "[t]ransparency is not just an end in itself, but an interim step on the road to intelligibility."[49]

Disclosure-to-consumer rules can require a user-friendly disclosure of various types of information,[50] including algorithms,[51]

---

[47] *See* LOUIS D. BRANDEIS, OTHER PEOPLE'S MONEY AND HOW THE BANKERS USE IT 92 (1914) ("Sunlight is said to be the best of disinfectants."); *see also* Frank Pasquale, *Beyond Innovation and Competition: The Need for Qualified Transparency in Internet Intermediaries*, 104 NW. U. L. REV. 105, 160–61 (2010) (proposing transparency rules to address neutrality issues of internet service providers).

[48] Sara Gerke, *"Nutrition Facts Labels" for Artificial Intelligence/Machine Learning-Based Medical Devices—The Urgent Need for Labeling Standards*, 91 GEO. WASH. L. REV. 79, 145 (2023).

[49] PASQUALE, *supra* note 36, at 8.

[50] For disclosure suggestions, see, for example, Gerke, *supra* note 48, at 152–58.

[51] *See, e.g.*, Danielle Keats Citron, *Open Code Governance*, 2008 U. CHI. LEGAL F. 355, 358 (2008).

inputs and outputs,[52] explanations of the AI-made (or AI-involved) decisions,[53] and the use of AI products in a specific occasion.[54] However, disclosure rules have some inherent limits, such as its inherent conflicts with trade secrets protection, insufficient testing capabilities, randomness issues, and time-based limitations.[55] Policymakers, therefore, should tailor the disclosure-to-consumer rules with care because excessive disclosure to the public also may risk privacy harms,[56] national security,[57] as well as trade secrets leaks and potential free-riding from competitors.[58]

(b) AI Literacy

AI literacy programs are an essentially educational undertaking. It aims at enabling consumers to know and understand, apply and analyze,

---

[52] *See, e.g.*, Anupam Chander, *The Racist Algorithm?*, 115 MICH. L. REV. 1023, 1024 (2017); Nicola Lucchi, *ChatGPT: A Case Study on Copyright Challenges for Generative Artificial Intelligence Systems*, EUR. J. RISK REGUL. 1, 21 (2023).

[53] *See, e.g.*, Sandra Wachter, Brent Mittelstadt & Chris Russell, *Counterfactual Explanations Without Opening the Black Box: Automated Decisions and the GDPR*, 31 HARV. J.L. & TECH. 841 (2018).

[54] *See, e.g.*, Philipp Hacker, Andreas Engel & Marco Mauer, *Regulating ChatGPT and Other Large Generative AI Models*, *in* 2023 ACM CONF. ON FAIRNESS, ACCOUNTABILITY, & TRANSPARENCY 1112, 1119 (2023).

[55] *See, e.g.*, Joshua A. Kroll et al., *Accountable Algorithms*, 165 U. PA. L. REV. 633, 657–60 (2017); Stefan Larsson & Fredrik Heintz, *Transparency in Artificial Intelligence*, 9 INTERNET POL'Y REV. 1, 8 (2020).

[56] *See* Kroll et al., *supra* note 55, at 658; *see also* Roger Allan Ford & W. Nicholson Price II, *Privacy and Accountability in Black-Box Medicine*, 23 MICH. TELECOMM. & TECH. L. REV. 1 (2016).

[57] *See* PASQUALE, *supra* note 36, at 193.

[58] *See* Frank Pasquale, *Restoring Transparency to Automated Authority*, 9 J. ON TELECOMM. & HIGH TECH. L. 235 (2011).

evaluate and even create AI.[59] AI literacy has many functions, but here we focus on its under-estimated role in the AI market, arguing that AI literacy programs act profoundly in closing information asymmetry and invigorating AI market competition by breeding more *ethical* and *critical* consumers.

First, through enhancing AI literacy, AI consumers will be capable of evaluating competing AI products and services with ethical yardsticks.[60] More educated AI consumers will leave less ethical AI operators and switch to more responsible and trustworthy ones. Consequently, less ethical operators will lose consumers and be left at a disadvantage in the marketplace. In this way, the commercial interests of AI operators are squarely aligned with the ethical expectations of the general public. In this sense, what we need more is ethical consumers, instead of mere ethical machines—ethical customers can spawn ethical machines. Therefore, policymakers can adopt AI literacy schemes to shape consumers' expectations for AI products, in which way externalities (*e.g.*, environmental pollution) created by AI operators can be internalized in the decision-making process of purchase.

AI literacy programs should also be pivoted around cultivating critical consumers. Critical consumers are essential for market competition because they can make relatively wiser choices when deciding on AI products or services. Critical consumers can discern and evaluate the risks involved in AI products or services. If unsatisfied with the quality or price of the current AI operator, they will switch to its competitors. Their leaving may be followed by a group of users, which will

---

[59] *See generally* Davy Tsz Kit Ng et al., *Conceptualizing AI Literacy: An Exploratory Review*, 2 COMPUTS & EDUC. A.I. 1 (2021) (proposing four aspects of AI literacy: (1) know and understand, (2) use and apply, (3) evaluate and create, and (4) ethical evaluations).

[60] *See* Davy Tsz Kit Ng et al., *AI Literacy: Definition, Teaching, Evaluation and Ethical Issues*, 58 PROC. ASS'N FOR INFO. SCI. & TECH. 504, 507 (2021) ("[I]nstead of merely knowing how to use AI applications, learners should be inculcated with the underlying AI concepts for their future career, as well as the ethical concerns of AI applications to become a responsible citizen.").

undermine the competitiveness of the AI incumbents. Here, it is noteworthy that the requirements for consumers to be critical are not as harsh as for consumers to be rational. "Rationality" is a concept that has been used to indicate "a cognitively intensive, calculating process of maximization in the self-interest."[61] This requirement proves too hard to implement because, in most cases, "self-interest" is not mathematically calculatable, and thus, consumers' choices inevitably deviate from the over-idealistic goal of self-interest maximization.[62] However, the fact that consumers are not rational does not mean that they are not critical when making decisions and that the market is not competitive. The effectiveness of market competition never relies on the perfect rationality of all members of society.

Given its fundamental value in invigorating the AI market competition and furthering the public interest, AI literacy programs should be fully or partially sponsored by the government.

(c) Right-and-Responsibility

Right-and-responsibility legislations (R&R) are those laws that authorize private parties to exercise their rights to subject the counterparties to some obligations so that the obligors will be deterred from causing harm. Also discussed in Part III as a value-directed approach, R&R can empower consumers and invigorate market competition. First, R&R can address lock-in effects and reduce consumers' switching costs. A notable example is data portability rights. Policymakers and researchers expect data to be portable so that consumers can take their data and change the service providers more easily, in which way

[61] Vernon L. Smith, *Rational Choice: The Contrast Between Economics and Psychology*, 99 J. POL. ECON. 877, 878 (1991).

[62] *Id.*

users will not be locked in by the incumbent market player and the companies will compete in a more even way.[63]

Second, endowment of rights can also elevate public AI literacy and enhance the consumers' critical evaluation of competing AI products in the marketplace. On the one hand, people with rights will value their rights and have autonomy to decide for themselves.[64] Consumers will take it for granted and choose to exercise their rights when necessary. This is indeed the foundation for an efficient operation of market mechanism.[65] For instance, consumers with legally recognized data autonomy can decide whether or not to share their information for LLMs' training, fully according to their own privacy expectations.[66] In addition, potential litigation against a certain AI company as well as corresponding news reports that disclose its misconduct will inform more consumers in the marketplace, putting the company under huge market and reputation pressure to remedy the problem and protect people's rights.

For those heavily value-laden R&R rules, we suggest considering them in the second layer. Although these R&R rules have great potential to empower consumers, they act primarily as a baseline to safeguard

[63] *See, e.g.*, Christopher S. Yoo, *When Antitrust Met Facebook*, 19 GEO. MASON L. REV. 1147, 1154 (2012) (noting the argument made by some scholars that "the inability to move data from one social networking site to another can create a form of lock-in"); Aaron Perzanowski & Jason Schultz, *Digital Exhaustion*, 58 UCLA L. REV. 889, 900 n.50 (2011) ("Switching costs would be reduced further if consumers were assured data portability between platforms."); Jay P. Kesan, Carol M. Hayes & Masooda N. Bashir, *Information Privacy and Data Control in Cloud Computing: Consumers, Privacy Preferences, and Market Efficiency*, 70 WASH. & LEE L. REV. 341, 470 (2013) ("[D]ata mobility in the cloud would facilitate consumer participation and reduce transaction costs for consumers when moving from one provider to another.").

[64] *See generally* Amartya Sen, *Well-Being, Agency and Freedom: The Dewey Lectures 1984*, 82 J. PHIL. 169 (1985); Amartya Sen, *Human Rights and Capabilities*, 6 J. HUM. DEV. 151 (2005).

[65] *See generally* MILTON FRIEDMAN & ROSE FRIEDMAN, FREE TO CHOOSE: A PERSONAL STATEMENT (1980).

[66] *See* Zhang, *supra* note 12, at 149–51.

human rights and freedoms. For those R&R rules that act mainly as a value-neutral market catalyst (*e.g.*, data portability), we suggest policymakers craft and implement them in the first layer.

B. Competition Adjustment

Competition adjustment approaches are commonly used across various fields. They aim for adjusting the competitive dynamics among competitors to foster a level playing field and promote a healthy competition environment. Specifically, competition adjustment approaches can lower the market entry barriers and empower those AI startups by mandating the sharing of key inputs, mitigating network effects, and prohibiting potential anti-competitive conducts of incumbent AI companies. Ultimately, the AI marketplace will become more competitive; consumers will be able to have more alternatives, which can pressure those incumbents into aligning their commercial interests with public expectations.

Some frameworks, such as competition law, break-ups (also known as unbundling rules), and non-discrimination rules (also known as "vertical integration restriction rules"), are well-known to many jurisdictions and some researchers have proposed these rules to address AI issues.[67] Our prescription below mainly addresses the market failures caused by the insurmountable costs to access key inputs, such as models, data, users, and talent resources. Accordingly, we suggest four methods—although not exhaustively––namely open source, interoperability, data access, and free labor flow, that can address these concerns and restore the market to a competitive equilibrium.

(a) Open Source

Open-source AI models are those models with essential components publicly available to the public, such as source codes,

[67] *See e.g.*, Vipra & Korinek, *supra* note 39, at 30–37; Narechania & Sitaraman, *supra* note 39, at 44–46.

datasets, and weights.[68] Free and open access to key inputs of AI models enables second use and modification in the downstream applications at significantly lower costs.[69] With collective efforts by the community, open-source also eases efficient detection and correction of errors for the models.[70] Compared to closed models, open-source AI models have the potential to distribute the concentrated power to the downstream market players, breeding more innovative and diverse AI ecosystem, and enhancing transparency in AI systems.[71]

Despite their great potentials, policymakers and technical experts are ambivalent about open-source AI models.[72] This is because open-source models are vulnerable to malicious uses.[73] Misusers could easily disable the safety components and apply the models to fulfill detrimental tasks, such as generating voluminous disinformation, deepfakes, and manufacturing biological and chemical weapons.[74] This, if left

[68] *See* Yash Raj Shrestha, Georg von Krogh & Stefan Feuerriegel, *Building Open-Source AI*, 3 NATURE COMPUTATIONAL SCI. 908, 908 (2023).

[69] *Id.*

[70] *Id.*

[71] *See* Rishi Bommasani et al., *Considerations for Governing Open Foundation Models*, STANFORD HAI (Dec. 13, 2023), https://hai.stanford.edu/issue-brief-considerations-governing-open-foundation-models; Ben Brooks, *Open-Source AI Is Good for Us: But Current Regulatory Trends Threaten to Quash Transparency and Competition in AI*, IEEE SPECTRUM (Feb. 8, 2024), https://spectrum.ieee.org/open-source-ai-good.

[72] *See e.g.*, EU AI Act, *supra* note 5, at art. 2 (12); Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence, No. 14,110, 88 Fed. Reg. 75,191, § 4.6 (Nov. 1, 2023) (to be codified at 3 C.F.R.), https://www.govinfo.gov/content/pkg/FR-2023-11-01/pdf/2023-24283.pdf.

[73] *See, e.g.*, Tom Simonite, *The AI Text Generator That's Too Dangerous to Make Public*, WIRED (Feb. 14, 2019), https://www.wired.com/story/ai-text-generator-too-dangerous-to-make-public/; 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.

[74] *See* Harris, *supra* note 73.

unregulated, will threaten and even undermine fundamental human values and freedoms.

As we will illustrate in Part III, market competition, understandably, cannot address intentional abuse of open-source AI models. Therefore, it is essential to set appropriate value-directed guardrails to prevent such misuse. However, that does not negate the role of open-source AI models in incubating diverse applications and catalyzing AI marketplace. Policymakers wanting a more competitive AI market should embrace open-source AI models, creating a lenient environment for wide experiments and applications. Policymakers should also tailor their regulatory tools to be both robust and proportionate, on the one hand, safeguarding human values and, on the other hand, minimizing chilling effects caused by unnecessary interventions and preserving the openness culture.[75]

(b) Interoperability

Interoperability is another approach to adjusting the competitive relationship between those technology incumbents and other smaller companies. It is broadly defined as "the ability to transfer and render useful data and other information across systems, applications, or components."[76] Interoperability package can include various instruments, such as real-time data sharing, data portability, intellectual property sharing, and non-discrimination requirements.[77]

[75] *See* Brooks, *supra* note 71.

[76] JOHN PALFREY & URS GASSER, INTEROP: THE PROMISE AND PERILS OF HIGHLY INTERCONNECTED SYSTEMS 5 (2012).

[77] *See* Herbert Hovenkamp, *Antitrust Interoperability Remedies*, 123 COLUM. L. REV. 1, 4 (2023).

Interoperability is especially effective when the market is highly concentrated due to economies of scale and network effects.[78] Without interoperability, T-Mobile subscribers cannot message Vodafone subscribers; Bank of America users cannot transfer money to Citigroup users. Eventually, most users will choose the one with the largest network for the highest efficiency, which would render the market tip to a single winner. That is the so-called "winner-take-all." [79] Conversely, with interoperability, users of different networks are in a larger realm where they can "communicate" seamlessly with each other, without being subject to fragmented systems. In this way, consumers can select networks fully based on their preferences, rather than network size. They can stay with their desired features, without losing connectivity with other networks. This allows of a competitive market where small-scale but innovative companies can thrive by serving distinct customer bases.[80]

Realizing that the layers of AI supply chains are prone to be concentrated, researchers and policymakers have proposed interoperability to address the potential market failure.[81] For example, policymakers may consider mandating interoperable application programming interfaces (APIs)—"sets of protocols which define how software components communicate with on another"—across cloud and model layers to ease application developers' data sharing and switching.[82] In the same vein, interoperability at the application layer can facilitate users' data porting from one provider to another.

---

[78] *See* Herbert Hovenkamp, *Antitrust and Platform Monopoly*, 130 YALE L.J. 1952, 1975 (2021).

[79] *See id.* at 1969–70, n.67.

[80] For more benefits of interoperability, see generally PALFREY & GASSER, *supra* note 76.

[81] *See* Joint Statement, *supra* note 41, at 2; *see also* Narechania & Sitaraman, *supra* note 39, at 46–47.

[82] Oscar Borgogno & Giuseppe Colangelo, *Data Sharing and Interoperability: Fostering Innovation and Competition through APIs*, 35 COMPUT. L. & SEC. REV. 1, 3 (2019); *see also* Narechania & Sitaraman, *supra* note 39, at 47.

However, two concerns are noteworthy here. First, policymakers should harmonize the interoperability rules with established competition (or antitrust) law.[83] For example, compulsory interoperability might come against the legally acknowledged freedom of private companies to refuse to deal.[84] It may also facilitate homogenization and free-riding, which may distort competitive order and harm consumers.[85] Second, policymakers should be aware of the privacy and security risks inherent in a wider interoperability context.[86] In this sense, Palfrey and Gasser perspicaciously elucidated that "an optimal, not a maximum, degree of interop [is desirable]" and "[t]here is no single form or optimal amount of interoperability that will suit every circumstance."[87] Policymakers should tailor the degree of interoperability to reflect their jurisdiction-specific conditions and fulfill their specific regulatory objectives.

(c) Data Access[88]

Data for the innovation in the digital ecosystem is much like the plankton for ocean life in the sea.[89] For AI companies, "[t]he more data, the better the program."[90] However, small companies often find it

[83] *See* Hovenkamp, *supra* note 77, at 23–34 (harmonizing interoperability rules with competition (antitrust) law).

[84] *Id.* at 26 ("Compelled interoperability could be an effective remedy, but it must be legislated, or else duty-to-deal doctrine would have to be expanded.").

[85] *Id.* at 23.

[86] *See* PALFREY & GASSER, *supra* note 76, at 75–88.

[87] *Id.* at 76.

[88] Free data access can—but not necessarily—be one part of an open-source license. Also, data sharing is an essential element of interoperability arrangements. *See generally* Borgogno & Colangelo, *supra* note 82. Here, we separate it out to highlight the data as another essential input for AI training.

[89] *See* VIKTOR MAYER-SCHÖNBERGER & THOMAS RAMGE, ACCESS RULES: FREEING DATA FROM BIG TECH FOR A BETTER FUTURE, ch. 4 (2022).

[90] Sara Brown, *Machine Learning Explained*, MIT SLOAN SCH. MGMT. (Apr. 21, 2021), https://mitsloan.mit.edu/ideas-made-to-matter/machine-learning-explained.

challenging to access sufficient data to elevate their services and compete effectively against those incumbents. Those tech giants, knowing the significant competitive advantages brought about by big data, have sufficient incentives to keep data exclusive.[91] Some legal frameworks also permit such conservative data practices, which further entrenches the data concentration.[92]

Policymakers aspiring to free data from dominant technology companies should consider imposing data-sharing duties on those data monopolists, so that those small companies can access sufficient data to train their models. Policy design entails answering the following questions: On the doctrinal aspect, first, what constitutes an appropriate legal basis for data access rules—competition law or ex-ante intervention (*e.g.*, essential facility doctrines)?[93] A derivative question is: under what conditions should a company be subject to data access duties? And, how to harmonize data access rules with privacy and security protection requirements?[94] More granular practical questions include: should data access follow an on-request, one-to-one contract-based sharing approach, or should it be made a common database that is fully publicly available to all entities? [95] How to design the data access scheme to incentivize

[91] *See* MAYER-SCHÖNBERGER & RAMGE, *supra* note 89, at 108–09.

[92] *See generally* Michal S. Gal & Oshrit Aviv, *The Competitive Effects of the GDPR*, 16 J. COMPETITION L. & ECON. 349 (2020).

[93] *Compare* Hovenkamp, *supra* note 78, at 2035 (viewing data pooling as an antitrust remedy) *with* Zachary Abrahamson, *Essential Data*, 124 YALE L.J. 867 (2014) (proposing to treat data fulfilling certain requirements as essential facility) *and* Inge Graef, Data as Essential Facility: Competition and Innovation on Online Platforms (2016) (Ph.D. dissertation, KU Leuven).

[94] *See* MAYER-SCHÖNBERGER & RAMGE, *supra* note 89, ch. 6.

[95] *Id.*

companies to keep collecting, analyzing, and sharing data given the common free-riding problem of using public goods.[96]

(d) Free Labor Flow

Talent resources are also key to an AI company's survival and competitiveness.[97] Free movement of labor forces can act as an engine to foster knowledge diffusion across the entire AI industry, thereby invigorating innovation and competition. However, just as dominant companies would keep data exclusive to their own use, they likewise have incentives to lock in AI talents to minimize market forces and protect their market positions. [98] Even worse, some of their talent management strategies are anticompetitive, but legally endorsed, such as signing the non-compete clause.[99]

Therefore, policymakers may also consider freeing labor flow to steer the talent market away from the concentration. Here, one promising template for other jurisdictions to reference is the U.S. Federal Trade Commission's (FTC) Non-Compete Clause Rule.[100] FTC systematically illustrates how non-competes may negatively affect competitive conditions in labor markets and product and service markets, based on solid empirical evidence.[101] It also sufficiently considers the asserted justification for non-competes—namely ensuring employers' incentives to

[96] *See* Michael Mattioli, *The Data-Pooling Problem*, 32 BERKELEY TECH. L.J. 179, 200–05 (2017).

[97] *See, e.g.*, *Generative AI Raises Competition Concerns*, FED. TRADE COMM'N (June 20, 2023), https://www.ftc.gov/policy/advocacy-research/tech-at-ftc/2023/06/generative-ai-raises-competition-concerns (noting that "[t]he talent companies can acquire and maintain may play a key role in not only the path, but also the rate, of generative AI's evolution").

[98] *Id.*

[99] *Id.*

[100] Non-Compete Clause Rule, 89 Fed. Reg. 38342 (May 7, 2024).

[101] *Id.*

invest in worker training and trade secret sharing among the workplace.[102] Otherwise, employers will reduce their investments as these investments are not legally and contractually protected.[103] Lastly, the FTC evaluates the effectiveness of comparable alternatives that are more narrowly tailored to realize the goals to protect the competitive conditions, such as using trade secret law, patent law, or signing non-disclosure agreements, invention assignment agreements, and fixed duration contracts.[104]

(e) Regulatory Sandbox

Regulatory sandbox originates in fintech regulation as a way to rebalance the need for financial stability and interest in endless fintech innovations. [105] The sandbox is designed as a "safe space" where businesses can experiment on their innovative fintech products supervised by regulators to ensure any involved risks are controlled to an acceptable level.[106] This regulatory idea has been transplanted to the field of AI regulation. Under the EU AI Act, AI regulatory sandbox is to provide "a controlled environment that facilitates the development, testing and validation of innovative AI systems for a limited time before their placement on the market or putting into service pursuant to a specific plan."[107]

In addition to its capabilities to enhance human values (discussed in Part III) and facilitate compliance (Part IV), regulatory sandbox also aims at restoring the competitive relationship to equilibrium. Research finds that small companies are typically left at a competitive disadvantage

---

[102] *Id.*

[103] *Id.*

[104] *Id.*

[105] *See generally* Dirk A. Zetzsche et al., *Regulating a Revolution: From Regulatory Sandboxes to Smart Regulation*, 23 FORDHAM J. CORP. & FIN. 31 (2017).

[106] *Id.* at 45.

[107] EU AI Act, *supra* note 5, at art. 53.

in market competition when facing stringent regulatory standards and heavy compliance costs.[108] Some would-be entrants that cannot afford the investment of time and money in compliance may quit the market, which will lead to fewer innovations, less competition, and greater market concentration.[109] With regulatory sandbox, those companies otherwise having difficulty complying with the standards can test their innovations in a protected environment where they do not need to bear heavy compliance costs. With guidance from regulators, regulatees can bring risks down to an acceptable level more easily and quickly.

However, regulatory sandbox design is more complicated than one might expect. Regulatory sandbox comprises a set of components, such as policy standard setting, entry and exit thresholds, working mode for both regulators and regulatees, and selection of regulatory tools inside the sandbox.[110] This entails a delicate balance and careful design of every single part to maximize the efficacy of the sandbox. When designing regulatory sandbox for AI companies, policymakers can gain some lessons by referencing the fintech sandbox of various jurisdictions.[111] A policy brief titled "*How Regulatory Sandbox Programs Can Promote Technological Innovation and Consumer Welfare*" compares different jurisdictions' sandbox programs and generates intensive policy suggestions on how to strategically design each sandbox element to

[108] *See* Jon Truby et al., *A Sandbox Approach to Regulating High-Risk Artificial Intelligence Applications*, 13 EUR. J. RISK REGUL. 270, 292 (2022) ("This allocation of risks and costs will likely result in the stifling of innovation, especially for SMEs that cannot afford the cost of the risks in a strict liability regime. SMEs may also face barriers to entry, for example, for procedures on full quality compliance.").

[109] *Id.*

[110] *See, e.g.*, Walter G. Johnson, *Caught in Quicksand? Compliance and Legitimacy Challenges in Using Regulatory Sandboxes to Manage Emerging Technologies*, 17 REGUL. & GOVERNANCE 709, 710, 712–16 (2023).

[111] *See* WORD BANK GROUP, GLOBAL EXPERIENCES FROM REGULATORY SANDBOXES (2020), https://documents1.worldbank.org/curated/en/912001605241080935/pdf/Global-Experiences-from-Regulatory-Sandboxes.pdf.

enhance consumer welfare.[112] Some AI-specific studies have also delved into this issue.[113]

Here, policymakers should particularly avoid the sandbox's side effects on the market. Regulatory sandbox is not a traditional market adjustment mechanism. It is not at all value-neutral throughout the whole process—the selection of eligible regulatees, testing and cooperation, and evaluation for exiting the sandbox are essentially involved with regulators' subjective judgment. This risks excessive market intervention and distortion of competitive equilibrium. One possibility is that sandbox participants may obtain some regulatory privilege over their competitors outside the sandbox.[114] In response, some researchers suggest that the lessons regulators obtain from the sandbox should be applied evenly to all similar market players, regardless of inside or outside the sandbox.[115] Another suggestion is that the benefits gained by sandbox participants should be time-limited, to avoid unnecessary government intervention.[116]

III. SECOND LAYER: VALUE-DIRECTED RULES

Market competition is not a panacea. Under specific conditions, market players, although under sufficient market pressure, may not keep improving the quality and lowering the price, but race to the bottom by adopting unethical means to maintain or improve their market positions.

---

[112] *See* Ryan Nabil, *How Regulatory Sandbox Programs Can Promote Technological Innovation and Consumer Welfare: Insights from Federal and State Experience*, COMPETITIVE ENTER. INST. 5–8 (Aug. 17, 2022), https://cei.org/wp-content/uploads/2022/08/Ryan_Nabil_-_Regulatory_Sandboxes-3.pdf.

[113] *See, e.g.*, Johnson, *supra* note 110, at 712–16; Ryan Nabil, *Artificial Intelligence Regulatory Sandboxes*, 19 J.L. ECON. & POL'Y 295, 327–43 (2024).

[114] *See generally* Brian R. Knight & Trace E. Mitchell, *The Sandbox Paradox: Balancing the Need to Facilitate Innovation with the Risk of Regulatory Privilege*, 72 S. CAL. L. REV. 445, 462–65 (2020); Nabil, *supra* note 113, at 325.

[115] *See, e.g.*, Nabil, *supra* note 113, at 329–30.

[116] *See, e.g.*, *id.* at 330; *see also* Knight & Mitchell, *supra* note 114, at 28.

Stucke and Ezrachi called this "toxic competition."[117] Some key market conditions may facilitate toxic competition, including but not limited to (1) that market player's commercial interests are not aligned with society's collective interests;[118] (2) that quality improvement or degradation is largely invisible to other market participants, including consumers and other competitors (i.e., aforementioned information asymmetry as market failure);[119] and (3) that market forces incentivize competitors to exploit consumer weaknesses and capitalize on consumers decision-making shortcomings.[120]

Toxic competition may also occur among AI companies. AI operators pursuing the maximization of commercial interests may brush aside some fundamental interests of consumers and adopt despicable strategies to exploit human weakness, which are largely invisible to and unrealized by consumers.[121] For example, AI operators may use AI to deliver biased and differentiated outcomes to harm consumer surplus when consumers are subject to insufficient information and imperfect rationality.[122] In addition, AI companies, wanting more personal data, may use lengthy, vague policy notices in overly technical wording to lure consumers into giving consent to undesired data processing, or set default designs and dark patterns that are unobtrusive to users but in favor of AI

---

[117] MAURICE E. STUCKE & ARIEL EZRACHI, COMPETITION OVERDOSE: HOW FREE MARKET MYTHOLOGY TRANSFORMED US FROM CITIZEN KINGS TO MARKET SERVANTS 40 (2020).

[118] *Id.* at 12–13.

[119] *Id.* at 63–64.

[120] *Id.* at pt. I, ch. 3.

[121] Geoffrey Hinton, the godfather of AI, has expressed such concerns in an interview after winning the Nobel Prize. *See* BBC Newsnight, *'Godfather of AI' on AI "exceeding human intelligence" and it "trying to take over"*, YOUTUBE (Oct. 8, 2024), https://www.youtube.com/watch?v=MGJpR591oaM ("[T]he competition is going to cause [AI] to be developed rapidly and the competition means that they won't put enough effort into safety).

[122] *See generally* Oren Bar-Gill, Cass R. Sunstein & Inbal Talgam-Cohen, *Algorithmic Harm in Consumer Markets*, 15 J. LEGAL ANALYSIS 1 (2023).

giants.[123] In such cases, consumers' expectations are not well aligned with AI operators' commercial interests. Of course, we can design our governing strategies to reconcile their divergent interests, by, for example, AI literacy or disclosure rules, but that is not enough. Without guardrails in place, competition may undesirably incentivize AI companies to race to the bottom. That is why sometimes the market mechanism is far from enough to make policymakers feel assured to relax further regulations.

Therefore, some guardrails should be set beforehand as the bottom-line rules to uphold human values and deter the would-be violators. These value-directed rules are necessary especially when some specific risks are unacceptable, no matter whether the market can work effectively. We place value-directed rules at the second layer because it would be impossible to delineate an accurate redline that AI operators and users must comply with in the absence of knowing the effectiveness of the market workings. Value-directed rules are not agile enough to respond to changing circumstances and unpredictable risks. If tailored inaccurate and inappropriate, these hard rules will unavoidably lead to false positives and false negatives.[124] This will distort market structure and undermine technological innovation in the long term.[125]

Thus, market mechanism and bottom-line rules are not mutually exclusive. Just as market failures justify some government intervention, these government failures necessitate the market mechanism that is comparatively more agile, efficient, neutral, and sustainable. Sinclaire perceptively opposed the flawed dichotomy between self-regulation and government intervention, arguing that "some combination of the two will

---

[123] *See* Ari Ezra Waldman, *Cognitive Biases, Dark Patterns, and the "Privacy Paradox"* 31 CURRENT OPINION IN PSYCH. 105, 105 (2020); WOODROW HARTZOG, PRIVACY'S BLUEPRINT 2 (2018).

[124] *See* BALDWIN, CAVE & LODGE, *supra* note 20, at 288–91.

[125] *See* Mark A. Lemley, *The Contradictions of Platform Regulation*, 1 J. FREE SPEECH L. 303, 330–35 (2021) (explaining why regulation is ineffective and even harmful to market competition and innovation).

be necessary to deliver optimal regulatory outcomes."[126] Indeed, the relationship between market-based self-regulation (1st layer) and external interventions (2nd & 3rd layers) is like a sliding scale—the more effective the market, the less necessary the value-directed rules and procedural controls, and vice versa. Policymakers adhering to this framework should analyze the effectiveness of the market, the risks, and the chosen regulatory tools iteratively and periodically. If market conditions change—for example, the market becomes more competitive and effective in mitigating risks—then regulators should accordingly lessen the second-layer regulatory tools.

Here, two toolkits we suggest are *command-and-control rules* and *right-and-responsibility*.

A. Command-and-Control

One form of such regulation is command-and-control (C&C). C&C rules denote a cluster of rules that "tak[e] [businesses'] discretion away by threatening sanctions aimed at bringing firms' interests into alignment with those of society, as a whole."[127] It includes various regulatory tools, like setting standards, imposing mandatory do's and don'ts, and establishing penalties for non-compliance.[128] The risks that fall below the bottom line of the standards are deemed per se unacceptable and unlawful.

The C&C rules are value-laden. They have been widely proposed and employed in various fields, particularly the environment, natural resources, food production, healthcare, transportation, finance, telecommunications, and other cutting-edge technologies, which are

[126] Darren Sinclair, *Self-Regulation Versus Command and Control? Beyond False Dichotomies* 19 L. & POL'Y 529, 552 (1997).

[127] Cary Coglianese & Evan Mendelson, *Meta-Regulation and Self-Regulation*, *in* THE OXFORD HANDBOOK OF REGULATION 146 (2012).

[128] *See* BALDWIN, CAVE & LODGE, *supra* note 20, at 107.

closely related to safety, health, freedom, privacy, security, and other fundamental human values. Generally, policymakers would use more and harsher C&C rules in the context of market failures relative to a competitive and effective market, as monopolists under few market forces are less likely to sacrifice their own commercial interests in exchange for enhancing human values and public interest.

Government intervention through C&C rules, like the market mechanism, is also subject to some failures.[129] First, C&C rules are not efficient and difficult for regulators to enforce. This is primarily because *ex-ante* standard-setting and *ex-post* enforcement need detailed information to ensure an appropriate and effective governance outcome. However, it proves time-consuming and costly for governments to fulfill this requirement.[130] Conversely, it is those decentralized companies that possess firsthand information and can act efficiently and flexibly.[131] And, due to this information asymmetry, regulators are prone to regulatees' capture, which will misguide the policymaking and render the rules deviated from maximization of public interest.[132] Second, those preset rules are not agile and flexible enough to accommodate all needs of regulatees. This may cause an excessive burden to some entities and raise market entry barriers to would-be entrants, thus favoring market incumbents and enhancing market concentration.[133] Third, C&C rules

[129] *See* Cass R. Sunstein, *Paradoxes of the Regulatory State*, 57 U. CHI. L. REV. 407 (1990).

[130] *See* JANIS D. BERNSTEIN, ALTERNATIVE APPROACHES TO POLLUTION CONTROL AND WASTE MANAGEMENT: REGULATORY AND ECONOMIC INSTRUMENTS 2 (1993) ("[Command-and-control strategies] are inefficient for the regulatory agency, which must have detailed information concerning production processes and the suitability of various pollution control devices. With diverse industries, it is extremely expensive and time-consuming to obtain the necessary information and expertise on each industry.").

[131] *See generally* Hayek, *supra* note 23.

[132] *See* Jean-Jacques Laffont & Jean Tirole, *The Politics of Government Decision-Making: A Theory of Regulatory Capture*, 106 Q.J. ECON. 1089 (1991).

[133] *See* Robert W. Hahn & Robert N. Stavins, *Incentive-Based Environmental Regulation: A New Era from an Old Idea*, 18 ECOLOGY L.Q. 1, 6 (1991) (noting that

may disincentivize companies to keep innovating. Some research found that once companies meet the pre-set environmental protection standards, they will no longer invest in pollution control technologies.[134] Fourth, C&C rules may cause long-term side effects. For instance, overreliance on C&C rules in natural resources management can deliver the outcome instantaneously but will lead to a fundamental loss of resilience in the natural ecosystem in the long run.[135]

Therefore, policymakers proposing C&C rules should realize these failures, tailoring the rules proportionately and harmonizing C&C tools with market-invigorating strategies and procedural controls. Some common C&C tools include standard setting, licensure, conduct-based duties, and prohibition and punishment. In this section, we introduce and evaluate the standard setting, licensure, and another novel tool—regulatory sandbox.

(a) Standard Setting

Standard setting constitutes the first step for C&C rules. In the risk management scenario, the established standard acts as a redline to prevent potential "race-to-the-bottom" that risks harming fundamental human values. Below the line are generally unacceptable risks that regulatees should take measures to avoid; above the line is the free arena where market players can compete by offering lower prices, higher qualities, or better services. The placement of this redline to a large extent reflects the regulatory goal, which is subject to historical, cultural, political, commercial, and various other factors specific to that

"uniform technology-based and performance standards . . . tend to lead to inefficient outcomes in which firms use unduly expensive means of controlling pollution"); *see also* Michal S. Gal & Oshrit Aviv, *The Competitive Effects of the GDPR*, 16 J. COMPETITION L. & ECON. 349 (2020) (explaining the anti-competitive effects caused by the GDPR).

[134] *See* BERNSTEIN, *supra* note 130, at 2.

[135] *See generally* C. S. Holling & Gary K. Meffe, *Command and Control and the Pathology of Natural Resource Management*, 10 CONSERVATION BIOLOGY 328 (1996).

jurisdiction.[136] However, drawing an appropriate redline is occasionally challenging for policymakers, as this process involves integrating various values some of which conflict with one another.[137]

In the context of AI governance, to establish appropriate standards for AI operators, policymakers should at least sequentially fulfill three tasks: (1) accurately identify *what* risks are involved in the AI value chain; (2) comprehensively evaluate *how* these risks impact human affairs; (3) proportionately draw a redline to delineate which kinds of AI risks are acceptable––and which are not. Indeed, the first two steps, namely risk identification and evaluation, should be conducted iteratively throughout the three layers.[138]

Researchers and policymakers have grasped the full picture of AI risks, but limited research focuses on how to evaluate the impacts of these risks at the methodological level. One impressive forthcoming article is titled "*Assessing the (Severity of) Impacts on Fundamental Rights*" authored by Malgieri and Santos.[139] In this research, Malgieri and Santos propose a "multi-metric approach" to measure and evaluate the severity of interferences to fundamental rights.[140] Specifically, they consider *normative objective parameters* measured by the cumulation of violation, duration, scope, and reversibility of the infringement; *normative subjective parameters* assessed by group's or individual's subjective perceptions; and *real-life tangible impacts on individuals* quantified by

[136] *See* BALDWIN, CAVE & LODGE, *supra* note 20, at 299.

[137] *See id.* at 292 (arguing that the difficulty in delivering regulatory undertakings stems from differing risk priorities among regulators, firms, and the public, and involves contentious political decisions on risk targeting, intervention levels, and acceptable risk thresholds); *see also* Howard Latint, *Good Science, Bad Regulation, and Toxic Risk Assessment*, 5 YALE J. ON REGUL. 89 (1988).

[138] *See* discussion *infra* Part V.

[139] *See* Gianclaudio Malgieri & Cristiana Santos, *Assessing the (Severity of) Impacts on Fundamental Rights*, (June 28, 2024), https://papers.ssrn.com/abstract=4875937.

[140] *Id.*

money, time, scale, well-being, and others.[141] Policymakers might find this framework helpful for evaluating the extent of AI risks to human affairs.

Recognizing all involved risks and their influences, policymakers should draw a proportionate redline to prevent those unacceptable risks. During this process, policymakers should realize that risk control per se comes with certain risks. Therefore, when delineating the redline, policymakers need to weigh the benefits of the proposed standards against their costs.[142] Here, it is noteworthy that a zero-risk approach to standard-setting will not always reduce risks to zero but may even cause unexpectedly severe side effects to the industry and society in the long term. For example, the Chinese government adopts a zero-risk approach by providing an exhaustive list of prohibitive content in its Article 4 of the Generative AI Interim Measure.[143] However, this provision is impossible to fulfill, as the working of generative AI is not based on deterrence from authorities, but on the statistical pattern—"the more frequently a claim appears in the dataset, the higher likelihood it."[144] If the whole information ecosystem is imperfect, then the input, and the output can by no means be zero-risk.[145] Additionally, AI-enabled chatbot is only one

[141] *Id.* at pt.3.

[142] *See* BALDWIN, CAVE & LODGE, *supra* note 20, at 298.

[143] *See* Shengcheng Shi Rengong Zhineng Fuwu Guanli Zanxing Banfa (生成式人工智能服务管理暂行办法) [Interim Measures for Regulating Generative AI Services] (promulgated by Cyberspace Admin., Nat'l Dev. & Reform Comm'n, Ministry Educ., Ministry Sci. & Tech., Ministry Indus. & Info. Tech., Ministry Pub. Sec., Nat'l Radio & Television Admin.), July 10, 2023, http://www.cac.gov.cn/2023-07/13/c_1690898327029107.htm. For an English-translated version, see, for example, Interim Measures for the Management of Generative Artificial Intelligence Services, CHINA L. TRANSLATE (July 13, 2023), https://www.chinalawtranslate.com/en/generative-ai-interim/.

[144] *See* Luke Munn, Liam Magee & Vanicka Arora, *Truth Machines: Synthesizing Veracity in AI Language Models*, AI & SOC. 1, 3 (2023).

[145] *See* Jiawei Zhang, *ChatGPT as the Marketplace of Ideas: Should Truth-Seeking Be the Goal of AI Content Governance?*, 35 STAN. L. & POL'Y REV. ONLINE 11, 25–35 (2024).

information source. It competes with various information outlets for quality, like speed, accuracy, and relevance.[146] To gain a higher market share, AI companies have endogenous incentives to derisk their systems and align their models with public expectations.[147] China's zero-risk standard may also create strong silencing effects to incumbent AI companies—they may train their models to become a "SorryGPT," from which consumers will get a lot of apologies but cannot receive their desired answers. This harsh approach will also deter the would-be market entrants, as fulfilling this extreme standard is very costly to a small or medium size company.[148]

(b) Licensure

Licensure is not foreign to policymakers. As an ex-ante regulatory approach, it requires product and service providers to comply with preset, mandatory standards in exchange for a license before entering the market. It means that without a license, provisions are deemed per se unlawful—– and therefore not allowed to approach the general public. The primary goal of licensure is to safeguard fundamental human values from potentially detrimental undertakings.[149] Under the licensure regime, the regulator acting as a gatekeeper has the final say on whether the present

---

[146] *See generally* Zhang, *supra* note 12.

[147] *Id.*

[148] *Id.*; *see also* Thibault Schrepel & Alex 'Sandy' Pentland, *Competition Between AI Foundation Models: Dynamics and Policy Recommendations* 16 (MIT Connection Science Working Paper, July 2024) (arguing that "rules and standards should not—unless strictly necessary—raise compliance costs to levels that small and medium-sized players cannot reasonably afford, force licensing, create unnecessary barriers to data access, etc.").

[149] *See, e.g.*, WHITE HOUSE, OCCUPATIONAL LICENSING: A FRAMEWORK FOR POLICYMAKERS 7 (2015), https://obamawhitehouse.archives.gov/sites/default/files/docs/licensing_report_final_nonembargo.pdf ("[L]icensing can benefit practitioners and consumers through improving quality and protecting public health and safety.").

company has fulfilled its standards and whether it is allowed to enter the market.

Given the opaque nature of AI systems and the unpredictable and irreversible harm AI may cause, Malgieri and Pasquale argue for adopting licensure in the governance of high-risk AI systems.[150] They argue that the ex-post explanation-consent approach cannot satisfactorily mitigate the risks.[151] Instead, licensure as a complement can ensure the accountability of AI companies and prevent those harms.[152] Undoubtedly, licensure can be a robust guardrail that can directly fence the market off any unwanted risks.

However, as a paternalistic approach, licensure can be problematic if it is not tailored proportionately. First, policymakers may encounter various challenges when designing licensure rules.[153] It is still unclear which kinds of AI systems should be subject to licensure, how to appropriately set the market-entry threshold and evaluation criteria, and how to update those criteria, keep tracking ex-post businesses and renew the license in such a fast-developing industry.[154] However, policymakers and regulators are not always well positioned to set these standards and criteria due to their lack of expertise and increasing workload to track all AI products and services.[155] If these standards are set disproportionate, the licensure system can raise market barriers, distort market competition, and ultimately harm consumer welfare.[156] In this sense,

[150] *See generally* Gianclaudio Malgieri & Frank Pasquale, *Licensing High-Risk Artificial Intelligence: Toward Ex Ante Justification for a Disruptive Technology*, 52 COMPUT. L. & SEC. REV. 105899 (2024).

[151] *Id.* at pt. 2.

[152] *Id.* at pt. 3.

[153] *See* Guha et al., *supra* note 42, at 6.

[154] *Id.* at 46–47.

[155] *Id.* at 47–49.

[156] *Id.* at 52–55.

policymakers should realize these side effects and very carefully tailor the licensure framework.

(c) Regulatory Sandbox

Regulatory sandbox can be seen as a variant of licensure. It softens the hard line of licensure by offering companies that otherwise may not be licensed to enter the market a safe place to test their innovations. But same as licensure, the regulatory sandbox is born as a value-enhancing tool—it aims to trade off the promotion of innovation against the protection of other values.[157] Regulators will specify and evaluate the preset goal of innovation promotion and value protection throughout the entire sandbox program.[158] Companies that fulfill the goals will be permitted to enter the market while those not aligned with the sandbox goal will be forced to exit the sandbox and automatically lose the privilege.[159]

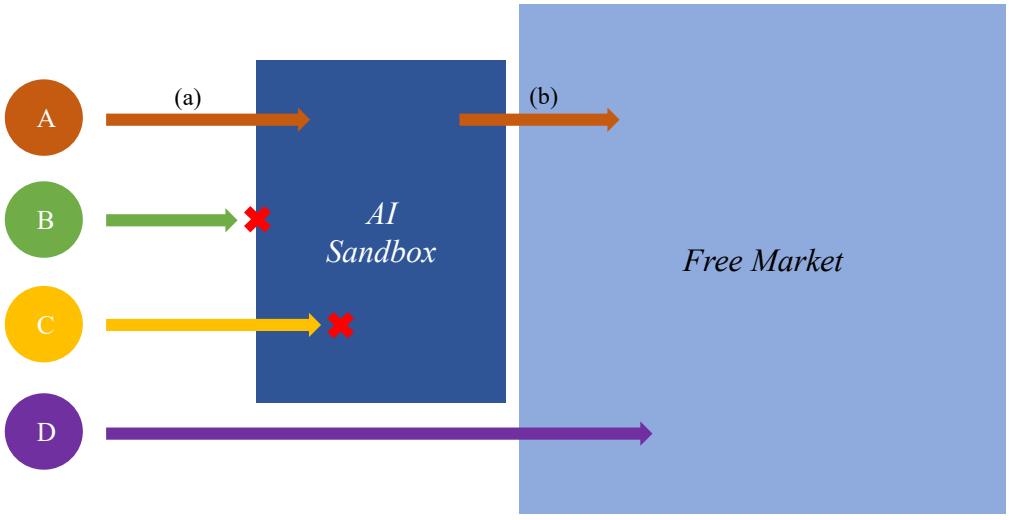
[157] *See* Hilary Allen, *Regulatory Sandboxes*, 87 GEO. WASH. L. REV. 579, 616 (2019).

[158] *See, e.g.*, Johnson, *supra* note 110, at 714–15.

[159] *See* Zetzsche et al., *supra* note 105, at 77.

Specifically, policymakers should conduct normative evaluations to determine which AI operators are eligible for entry into the sandbox (Cases A & C), which can enter the market without the sandbox (Case D), which are excluded from both the sandbox and the broader market (Case B), as well as which AI operator should be dropped out of the sandbox (Case C), and which can go to the market (Case A). Each decision should be guided by clear objectives and parameters to avoid capricious interference in the market. Therefore, policymakers should proportionately set goals for the decision process and the factors to be considered for entry and exit of the AI sandbox.

Decision Process throughout the AI Sandbox



In addition to the risks of excessive market intervention, policymakers should also be vigilant about the deceptive effectiveness of sandbox programs. For example, Omarova challenges the effectiveness of the regulatory sandbox by arguing that insights gained inside of the sandbox are not necessarily applicable in an uncontrolled public

market.[160] And, it is still unclear how to monitor those successful testees after they enter the public market (Case A, Process (b)).[161] In this way, there remains the risk that the harms that occur in the broader market are not observable in a limited, controlled setting. Therefore, policymakers aiming for robust governance must take such risks into consideration and conduct regular and iterative risk identification and evaluation.

B. Right-and-Responsibility

Another common form of regulation is rights and responsibilities (R&R). Compared to the C&C rules that are enforced by the government, the R&R rules are bottom-up and decentralized.[162] As discussed in Part II, the R&R approach empowers consumers and plays a vital role in the market mechanism. Here, we emphasize its value-enhancing function. As Joseph Raz put it, "the values underlying the (legally or politically) entrenched rights depend upon the rights, such that the (legal or political) entrenchment of the rights ultimately facilitates and contributes to the realization of the values."[163] AI policymakers can draw on R&R rules to safeguard human values, such as life, health, freedom, autonomy, equality, reputation, privacy, likeness, and creativity. R&R can act as a deterrent like C&C rules to those would-be violators, as the violation is always followed by litigation, compensation, and degraded reputation.

However, the R&R regime shares flaws similar to the C&C rules. First, harsh R&R rules, such as using strict liability, mean heavy compliance costs, which will lead to chilling effects that hamper

---

[160] *See* Saule T. Omarova, *Technology v Technocracy: Fintech as a Regulatory Challenge*, 6 J. FIN. REGUL. 75, 119 (2020) (arguing that "[t]he fundamental flaw of any regulatory sandbox, however, is the inherently limited capacity of 'controlled testing' of an individual fintech product to yield reliable insights about the systemic impact of that product outside the sandbox").

[161] *Id.*

[162] *See* Douglas A. Kysar, *The Public Life of Private Law: Tort Law as a Risk Regulation Mechanism*, 9 EUR. J. RISK REGUL. 48, 54 (2018).

[163] ALON HAREL, WHY LAW MATTERS 14 (2014).

innovation and competition.[164] Second, as the effectiveness of the C&C rules depends largely on enforcement, the deterring effect of the R&R regime is also subject to the capabilities of right holders to actively exercise their rights. If consumers have limited literacy to understand the harms, few access to know the infringement, and no incentives to file a lawsuit to protect their rights, then R&R rules are destined to fail. Additionally, the effectiveness of the R&R regime is also dependent on the proper level of compensation awarded by the courts, which, similar to the governments, frequently have difficulty accessing sufficient information to calculate a reasonable amount. [165] Therefore, legislators should enhance R&R rules with other regulatory tools, like disclosure rules and AI literacy schemes.

IV. THIRD LAYER: PROCEDURAL CONTROLS

As explained, government intervention—regardless market invigoration or value-directed rules—is often subject to certain failures. Thus, some reinforcement controls would be necessary. Here, we suggest procedural control as the third layer governance to guarantee the compliance of regulatees and the ultimate delivery of desirable regulatory outcomes envisaged in the first two layers.[166] Similar to audit regulation or "management-based regulation" raised by Coglianese and Lazer, the procedural control rules require companies to "produce plans that comply with general criteria designed to promote the targeted social goal," including "identification of hazards, risk mitigation actions, procedures

[164] *See generally* Truby et al., *supra* note 108; *see also* Miriam Buiten, Alexandre de Streel & Martin Peitz, *The Law and Economics of AI Liability*, 48 COMPUT. L. & SEC. REV. 105794 (2023).

[165] BALDWIN, CAVE & LODGE, *supra* note 20, at 127 ().

[166] *See* Cary Coglianese, *Compliance Is the Bridge to Better Regulatory Outcomes*, REGUL. REV. (Feb. 12, 2024), https://www.theregreview.org/2024/02/12/coglianese-compliance-is-the-bridge-to-better-regulatory-outcomes/ ("[J]ust adopting a regulation will not guarantee that regulatory problems are solved. Regulations will only work if they influence corporate behavior so that it aligns better with society's overall well-being.").

for monitoring and correcting problems, employee training policies, and measures for evaluating and refining the firm's management with respect to the stated social objective."[167] In addition, "firms are required to produce documentation of subsequent compliance or are subjected to reviews by regulators or third-party auditors to certify compliance."[168] Regulators should foster institutionalization of the risk identification system and event reporting procedures.[169] These approaches have been applied to regulate various fields, like food and industry safety.[170]

Here, two dimensions of procedural control rules can be distilled. First, it requires companies to follow some steps to control risks via their own internal management process. Second, it mandates specific information duties, such as documentation and disclosure of essential information, to enable the government to monitor and audit their compliance. These regulatory approaches are applicable in AI governance.

A. Internal Management

Internal management rules can be viewed as companies' self-regulation or internal audits, but it is different from the market mechanism. Under market forces, companies self-regulate to mitigate risks *automatically and voluntarily*, which is in alignment with their commercial interests. However, internal management as a procedural control approach establishes legal obligations that companies are *forced* to comply with, which might sometimes deviate from the maximization of companies' commercial interests. When compared to those value-based C&C rules, internal management is less value-laden. It proves more flexible by offering companies sufficient latitude in tailoring their

[167] Cary Coglianese & David Lazer, *Management-Based Regulation: Prescribing Private Management to Achieve Public Goals*, 37 L. & SOC'Y REV. 691, 694 (2003).

[168] *Id.*

[169] Charles Sabel, Gary Herrigel & Peer Hull Kristensen, *Regulation under Uncertainty: The Coevolution of Industry and Regulation* 12 REGUL. & GOVERNANCE 371 (2018).

[170] *See* Coglianese & Lazer, *supra* note 167, at 696–99.

strategies to comply with the rules.[171] For example, legislation can require AI companies to establish internal risk management mechanism to identify, rate, mitigate, and report risks. Without this rule, some companies may not invest in this system building as this may run counter to their commercial interests. Still, this rule does not encroach upon AI companies' autonomy—they retain substantial discretion to develop their own risk management system.

However, internal management rules are subject to specific limitations. For example, some researchers challenge its effectiveness in regulating low-probability, high-consequence events. [172] Some researchers argue that such internal regulation "simply relocate the problems (from outside to inside the firm), rather than solving them," supported by empirical case studies.[173] Given these shortcomings, we argue, again, that internal management rules and other regulatory toolboxes (*e.g.*, the market mechanism and command-and-control rules) are not mutually exclusive. Policymakers should adopt internal management rules concurrently with other regulatory tools to enhance the companies' compliance and to ensure a desirable outcome in a complicated setting.

In order to design and set up internal management rules, policymakers may consult a 2003 article titled "*Management-Based Regulation*" written by Coglianese and Lazer.[174] According to their proposal, regulators should investigate whether AI companies have

[171] *Id.* at 701–02.

[172] *See* Lori S. Bennear, *Beyond Belts and Suspenders Promoting Private Risk Management in Offshore Drilling* (Duke Environment Economics Working Paper No. EE 11-11, Dec. 2011), https://sites.nicholasinstitute.duke.edu/environmentaleconomics/wp-content/uploads/sites/3/2013/01/WP-EE-11-11.pdf.

[173] *See* Neil Gunningham & Darren Sinclair, *Organizational Trust and the Limits of Management-Based Regulation*, 43 L. & SOC'Y REV. 865, 896 (2009).

[174] *See* Coglianese & Lazer, *supra* note 167, at pt.III (designing management-based regulation).

incentives to *plan* and *implement that plan* on its own, to fulfill the regulatory goals.[175] If the answer is yes, then additional government intervention would be unnecessary.[176] Policymakers also need to craft their directing and monitoring strategies according to the AI companies' compliance incentives.[177] Noteworthily, some researchers have also developed an internal framework for AI companies to manage risks.[178]

B. Information Duty

(a) Documentation

Documentation is also an internal management approach that requires the regulatees to record and explain certain important matters. This constitutes a prerequisite for consistent and accurate information disclosure to consumers and to regulators. Consumers can gain knowledge of their targeted products or services from such disclosure and make wiser choices; regulators can review these prepared documents to audit regulatees' compliance states and decide whether to take further regulatory actions.

For instance, the EU's AI Act adopts documentation requirements to regulate high-risk AI systems. Article 11 requires that "[t]he technical documentation of a high-risk AI system shall be drawn up before that system is placed on the market or put into service and shall be kept up-to date."[179] Essential elements of the documentation include a description of the AI system, its elements and development process, the monitoring, functioning and control methods, its performance metrics, risk

[175] *Id.* at 13–18.

[176] *Id.*

[177] *Id.* at 19–23.

[178] *See generally* Jonas Schuett, *Three Lines of Defense against Risks from AI*, AI & SOC'Y (2023); Leonie Koessler, Jonas Schuett & Markus Anderljung, *Risk Thresholds for Frontier AI*, (June 20, 2024), http://arxiv.org/abs/2406.14713.

[179] EU AI Act, *supra* note 5, at art. 11.

management system, any changes made to the system throughout its lifecycle, compliance list and declaration, and post-market evaluation scheme.[180] Article 12, to further enhance traceability and accountability, mandates record-keeping duties, requiring high-risk AI systems to log the period of each use, reference database used, matched input data, and the people who conduct the human oversight.[181] Based on these rules, some articles delve into designing more detailed, implementable framework for complying with the documentation rules.[182]

Drawing on the EU's practices as well as existing research, worldwide policymakers should dissect the AI value chains and require a systematic documentation of the information beforehand that would be necessary for disclosure-to-consumer and disclosure-to-regulator.

(b) Disclosure-to-Regulator

Disclosure-to-regulator, unlike disclosure-to-consumer as a *bottom-up* paradigm (see Part II), acts more in assuring *centralized* regulators of firms' compliance. Regulators need sufficient, accurate information to audit companies' operations, spot and assess risks, (re)evaluate the effectiveness of various regulatory tools, and make necessary adjustments accordingly.[183] Regulators can also select some

[180] EU AI Act, *supra* note 5, at annex IV.

[181] EU AI Act, *supra* note 5, at art. 12.

[182] *See generally* Isabelle Hupont et al., *Documenting High-Risk AI: A European Regulatory Perspective*, 56 COMPUT. 18 (2023); Delaram Golpayegani et al., *AI Cards: Towards an Applied Framework for Machine-Readable AI and Risk Documentation Inspired by the EU AI Act* (2024), http://arxiv.org/abs/2406.18211; Bill Marino et al., *Compliance Cards: Computational Artifacts for Automated AI Regulation Compliance* (2024), https://arxiv.org/pdf/2406.14758.

[183] *See* Coglianese, Zeckhauser & Parson, *supra* note 34, at 278 ("[R]egulators need detailed and accurate information about the operations of private business enterprises to understand the scope and cause of regulatory problems, and to craft effective solutions to them"); Bradley C. Karkkainen, *Information as Environmental Regulation: TRI and Performance Benchmarking, Precursor to a New Paradigm?*, 89 GEO. L.J. 257, 263 (2001) ("A pervasive information bottleneck constrains the extent, effectiveness,

information and disclose it to the public so that informed consumers can make wiser choices, in which way disclosure-to-regulator acts indirectly to enhance the market mechanism.[184] Given these merits, such information-based regulation has long been proposed by researchers in various fields.[185]

However, disclosure-to-regulator rules are not always effective, as regulatees do not always have sufficient incentives to disclose sensitive information to help regulators spot their vulnerabilities and regulate accordingly.[186] Coglianese, Zeckhauser, and Parson are correct that "if competitors differ in the costs of controlling a certain type of risk, it may be beneficial for a low-cost firm to disclose information about the risk to the regulator."[187] Here, those AI giants have relatively more capabilities than those start-ups to bear the compliance costs, thus more likely and willing to disclose the risks. But it is still possible for those dominant AI companies to take collective silence tactic to blind policymakers. An interesting observation is their attitude towards copyright risk. AI giants' technical reports and model cards cover various—almost exhaustive—

efficiency, and responsiveness of the regulatory system."); Paul R. Kleindorfer & Eric W. Orts, *Informational Regulation of Environmental Risks*, 18 RISK ANALYSIS 155, 156 (1998) (arguing that the disclosed information closes the information gap between firms and the government, enabling regulators to examine the level of the risks at different stages of regulation and better tailor their policies and enforcement).

[184] *See, e.g.*, Cass R. Sunstein, *Informational Regulation and Informational Standing: Akins and Beyond*, 147 U. PA. L. REV. 613, 619 (1999) (arguing that some forms of disclosure and production of information "are designed to assist consumers in making informed choices; such statutes are meant to be market-enhancing.").

[185] *See generally* William M. Sage, *Regulating through Information: Disclosure Laws and American Health Care*, 99 COLUM. L. REV. 1701 (1999); Cass R. Sunstein, *Informing America: Risk, Disclosure, and the First Amendment*, 20 FLA. ST. U.L. REV. 653 (1993); Kleindorfer & Orts, *supra* note 183.

[186] *See* Coglianese, Zeckhauser & Parson, *supra* note 34, at 290–92.

[187] *Id.* at 292.

risks, but strategically exclude copyright risk.[188] However, given that several copyright lawsuits against AI companies are ongoing, copyright infringement is obviously a concrete risk inherent in LLM training. Their intention is evident: the companies do not want to inform regulators and the public of the copyright issues, as they believe it is impossible or at least devastatingly costly to clear copyrighted materials from its datasets.[189]

In response to potential collective silence, some regulatory strategies can be necessary. Two articles are quite instrumental in crafting disclosure rules, one general research titled "*Seeking Truth for Power*," suggesting strategies and tactics that regulators can use to secure information from regulates,[190] and the other AI-specific research titled "*Private Accountability in the Age of AI*," discussing regulatory methods to enhance AI transparency.[191] Here, we distill and integrate some useful key strategies for AI policymakers.

First, regulators can access information from the target's competitors. They can draw upon the asymmetries of interest among different market players to access information that is applicable to the whole industry from those that are prone to cooperate with regulators.[192] Regulators can also acquire information from the target's employees,

---

[188] *See* OpenAI, *GPT-4 Technical Report* (Dec. 19, 2023), https://arxiv.org/abs/2303.08774v4; Google, *Gemini: A Family of Highly Capable Multimodal Models* (June 17, 2024), https://arxiv.org/abs/2312.11805; Anthropic, *The Claude 3 Model Family: Opus, Sonnet, Haiku* (2024), https://www-cdn.anthropic.com/de8ba9b01c9ab7cbabf5c33b80b7bbc618857627/Model_Card_Claude_3.pdf.

[189] *See* James Titcomb & James Warrington, *OpenAI Warns Copyright Crackdown Could Doom ChatGPT*, THE TELEGRAPH (Jan. 7, 2024), https://www.telegraph.co.uk/business/2024/01/07/openai-warns-copyright-crackdown-could-doom-chatgpt/.

[190] *See* Coglianese, Zeckhauser & Parson, *supra* note 34.

[191] *See generally* Sonia K. Katyal, *Private Accountability in the Age of Artificial Intelligence*, 66 U.C.L.A. L. REV. 54 (2019).

[192] *See* Coglianese, Zeckhauser & Parson, *supra* note 34, at 297–99.

namely whistle-blowers.[193] Whistleblowing is particularly suitable in AI governance because, as argued by Katyal, it balances the need to enhance the transparency of those powerful but opaque AI systems as well as their potential misconducts[194] and the need to protect companies' interests in their trade secrets from excessive disclosure. [195] When securing information from market players and internal whistleblowers, regulators should adequately protect the sources of information.[196] Second, when collective silence occurs, regulators can create new incentives to nudge otherwise silent AI companies to disclose the needed information. [197] Third, regulators may design and utilize various channels to access their desired information, such as conditioning market-entry licenses on information release, mandating or rewarding disclosure, and formal (*e.g.*, public hearing) or informal interaction (*e.g.*, private communication) with the regulatees.[198] Here, AI policymakers should know the strengths and weaknesses of each channel and integrate them for better efficacy.

V. RISK ASSESSMENT AND REGULATION ADJUSTMENTS

Risk assessment, as a precondition for designing regulatory methods to control recognized risks and promote the public interest,[199] should be conducted throughout the entire process of policymaking and

[193] *Id.* at 299–300.

[194] *See* Katyal, *supra* note 191, at 126–29.

[195] *Id.* at 129–37.

[196] *See* Coglianese, Zeckhauser & Parson, *supra* note 34, at 299–300; *see also id.* at 131.

[197] *See* Coglianese, Zeckhauser & Parson, *supra* note 34, at 300–02.

[198] *Id.* at pt.III.

[199] *See* Julia Black & Robert Baldwin, *Really Responsive Risk-Based Regulation* 32 L. & POL'Y 181, 184 (2010); Sabel, Herrigel & Kristensen, *supra* note 169, at 372 ("The regulator's task was to elicit from firms the information necessary to establish public-regarding but economically feasible standards and rules, while avoiding "capture" or ceding regulatory control to its addressee.").

even the ex-post evaluations of the established regulatory arrangement. For a long time, the established risk-based regimes have been criticized as an institutionalized process that fails to recognize and manage emerging risks.[200] This is essentially because the risk is not static, but dynamic, changing with various forces, including but not limited to the political climate, current state of the art, market effectiveness, legal coercion, public mindsets, and literacy level. Therefore, policymakers should consider the foregoing factors and address the following risk-related questions throughout their relevant policymaking process:[201]

> (1) What risks have arisen from the new situation? (*risk identification*)
>
> (2) Which risks are per se acceptable, to what extent acceptable, and which risks are unacceptable? (*risk evaluation*)
>
> (3) Can the established or proposed regulatory tools control the identified risks to a reasonable level? (*risk evaluation → regulation evaluation*)
>
> (4) If yes to (3), are there any tools unnecessary and that can be dumped? (*regulation evaluation*), and
>
> (5) If no to (3), what other methods can be adopted to further mitigate the risks? (*regulation design*)

Examination of these questions should be conducted on a regular basis, which can inform policymaking and adjustments, ensuring a timely and proportionate regulatory response to a new situation. Meanwhile, policymakers should ensure that the adaptive policy adjustments have a

[200] *See* BALDWIN, CAVE & LODGE, *supra* note 20, at 288.

[201] *Id.*; William W. Lowrance, *The Nature of Risk*, *in* SOCIETAL RISK ASSESSMENT: HOW SAFE IS SAFE ENOUGH? 10-13 (Richard C. Schwing & Walter A Albers eds., 1980).

sound legal basis and do not cause fundamental uncertainties to the legal order.

To that end, setting "sunset provisions" for some first- and second-layer rules can be very helpful. Sunset provisions provide for a termination date of regulatory rules and require re-evaluation and sufficient justifications for the extension of their application period.[202] Sunset provisions have been widely used in various fields to assess the present regulatory effectiveness and efficiency and to deter unnecessary bureaucratic growth.[203] For example, the FCC used the sunset provision to regulate the fast-evolving cable industry. In 1992, the FCC was directed to enact prohibitions on exclusive contracts for cable and broadcast programming between a cable operator and a cable programming vendor, if the cable operator holds a stake in the vendor.[204] However, this exclusivity prohibition provision was subject to a sunset provision, which provides that such prohibition "shall cease to be effective 10 years after October 5, 1992, unless the Commission finds, . . . that such prohibition continues to be necessary to preserve and protect competition and diversity in the distribution of video programming."[205] Although this exclusivity prohibition provision survives ten years longer than the initially set duration due to the finding of the necessity to protect the market,[206] the sunset provision proves effective in curbing unnecessary government interference and regulatory expansion in the telecommunications market.

Therefore, we suggest that policymakers should render some first- and second-layer regulatory tools be subject to sunset provisions. According to sunset provisions, AI regulators should conduct periodical

---

[202] *See* Dan R. Price, *Sunset Legislation in the United States*, 30 BAYLOR L. REV. 401, 401 (1978).

[203] *Id.* at 417.

[204] 47 U.S.C. § 548(c)(2)(D).

[205] 47 U.S.C. § 548(c)(5).

[206] Cablevision v. FCC, 597 F.3d 1306 (D.C. Cir. 2010).

assessments of the necessity and effectiveness of their regulatory rules and terminate those outdated and inappropriate at due time. This ensures that the regulatory process is sufficiently agile and flexible in the face of rapidly changing circumstances led by AI and other emerging technologies.

VI. TOWARDS AGILE, ROBUST, AND PROPORTIONATE AI GOVERNANCE

Agility, robustness, and proportionality are three core tenets of AI governance. This is respectively due to three palpable realities of AI technologies. First, AI, as well as other emerging technologies, is developing much faster than the law. Second, AI has great potential to threaten people's ability to safeguard our collective values and public goods. Third, AI and other technological advancements can also significantly enhance human welfare and societal progress. Our design of the three-layered framework, sufficiently considering these realities, will embody and realize AI governance with agility, robustness, and proportionality.

First, the three-layered framework is aligned with the goal of agile AI governance. Agile governance is especially suitable for rapidly advancing and unpredictable technological landscapes. [207] Agile AI governance requires the entire AI governance framework to be highly

[207] *See* Alan F. T. Winfield & Marina Jirotka, *Ethical Governance Is Essential to Building Trust in Robotics and Artificial Intelligence Systems*, 376 PHIL. TRANSACTIONS ROYAL SOC'Y A 1, 2 (2018); World Economic Forum, *Agile Governance: Reimagining Policy-Making in the Fourth Industrial Revolution*, (Jan. 2018), https://www3.weforum.org/docs/WEF_Agile_Governance_Reimagining_Policy-making_4IR_report.pdf; OECD, *Practical Guidance on Agile Regulatory Governance to Harness Innovation* (2020), https://www.oecd-ilibrary.org/governance/reviewing-the-stock-of-regulation_1a8f33bc-en; Wendell Wallach & Gary E. Marchant, *An Agile Ethical/Legal Model for the International and National Governance of AI and Robotics*, https://www.aies-conference.com/2018/contents/papers/main/AIES_2018_paper_77.pdf.

flexible and promptly responsive to evolving technologies and the associated risks.[208] Essentially, agility is highly dependent on timely information collection and sophisticated information processing. As we explained, the market players are better positioned than central regulators in gathering and processing the most cutting-edge and first-hand information.[209] Policymakers' role, therefore, should pivot around timely adjusting rules to better respond to the realities, maximizing market functions, and minimizing unnecessary onerous duties. The three-layered framework equips policymakers with an operable guide to establish a flexible and adaptive regulatory framework, learn relevant expertise, and collaborate with scientists and industry experts so that they can possess the necessary capabilities to (re)assess the involved risks and (re)adjust the regulatory schemes.[210] These approaches can ensure a quicker response to the fast-changing realities, avoid the mismatch between old-fashioned rules and new facts, and promote sustained innovation and technological advancement.

Second, the three-layered framework can ensure robust AI governance. Robustness of governance indicates "the ability to cope with a variety of failure-inducing circumstances and behavior."[211] When regulating AI, potential market failures necessitate sufficient and effective rules and implementations to control the identified AI risks to an acceptable level so that fundamental human values can be upheld

---

[208] *See* National Academy of Public Administration & Project Management Institute, *Agile Regulation: Gateway to the Future*, at 7 (June 2022), https://s3.us-west-2.amazonaws.com/napa-2021/Agile-Regulation-Gateway-to-the-Future-Report.pdf.

[209] *See supra* note 23 and accompanying text.

[210] *See* Cary Coglianese, *Regulating Machine Learning: The Challenge of Heterogeneity*, COMPETITION POL'Y INT'L 8; Cary Coglianese, *Evaluating Regulatory Performance*, 8 U. PA. J.L. & PUB. AFFS. 47, 117 (2024) (noting that "the introduction of new technologies and their associated uncertainties of potential harm places a premium on ongoing efforts to gather information and maintain regulatory vigilance and agility").

[211] Patrick Honohan & Joseph E. Stiglitz, *Robust Financial Restraint*, *in* FINANCIAL LIBERALIZATION: HOW FAR, HOW FAST? 31 (Gerard Caprio et al., eds., 2001).

regardless of evolving realities.[212] To that end, policymakers guided by the three-layered framework should draw an appropriate redline to proscribe anticompetitive conduct and address potential toxic competition (or race-to-the-bottom) issues in which fundamental human values are in peril. In addition, the third layer, the procedural control rules, requires policymakers to craft specific procedural control rules to audit regulatees' businesses and ensure their compliance. Our prescription also warns that robustness should not be regarded as exhaustiveness—it is impossible, unnecessary, and even detrimental to cover all factors in policymaking and regulate AI and cutting-edge technologies in a zero-risk way.[213] In fact, the exhaustiveness approach is against the ethos of agility and proportionality.

Third, the three-layered framework can better generate proportionate AI governance. Proportionate AI governance means that regulatory instruments are prudently selected and regulatory severity is carefully calibrated to avoid chilling effects on AI companies. The proportionality test, composed of four elements—proper purpose, rational connection, necessity, and proportionality *stricto sensu* (balancing)—has been widely used in policymaking and judicial rulings to determine the legality of government actions.[214] Here, the proportionality test in AI policymaking primarily accentuates the last two elements,

---

[212] *See* CHRISTOPHER ANSELL ET AL., ROBUST GOVERNANCE IN TURBULENT TIMES 32 (2024) ("[R]obustness can be defined as the ability of a particular unit to continue to uphold some core functions, purposes, and values and/or maintain key structural or operational architectures in the face of disruptive perturbations by means of adaptation and innovation").

[213] *See* Lori S. Bennear & Jonathan B. Wiener, *Adaptive Regulation: Instrument Choice for Policy Learning over Time* (manuscript at 12) (Feb. 12, 2019) (arguing that "attempts to achieve a comprehensiveness [is] beyond human capacity"); *see also* Zhang, *supra* note 145 (arguing that zero-risk policy in AI content governance is impossible); Tonja Jacobi & Matthew Sag, *We Are the AI Problem*, 74 EMORY L.J. ONLINE 1 (2024) (showing the technical and societal challenges to de-risk generative AI models).

[214] *See* AHARON BARAK, PROPORTIONALITY: CONSTITUTIONAL RIGHTS AND THEIR LIMITATIONS 131 (2012).

namely necessity and balancing. Unlike the *Nirvana* approach, the three-layered framework informs policymakers of both the benefits and costs of every single regulatory tool. For example, imposing onerous and harsh AI regulation may address present issues promptly and directly, but can also cause long-term side effects, raising market barriers and deterring competition and further innovations.[215] Ultimately, consumer welfare will be worse off. Even worse, monopolistic incumbents under less market forces will eventually trigger harsher regulation, which is actually a vicious circle.[216] Policymakers following our proposal can avoid such onerous and excessive policies, as the three-layered framework is centered around setting achievable regulatory objectives, conducting meticulous and periodic risk assessments, selecting and harmonizing optimal regulatory tools, and reevaluating and updating the entire governance system. AI governance in this way can fully unlock the potential of AI to enhance human welfare.

Predictably, AI governance with sufficient agility, robustness, and proportionality will also safeguard downstream substantive values––including but not limited to democracy, social equity, safety, technological innovation, economic prosperity, and protection of fundamental human rights and freedoms––and contribute to achieve a fair balance among

[215] *See* Guha et al., *supra* note 42, at 6 ("[S]ome regulation proposals could—even if potentially useful in advancing legitimate public objectives—function to advantage powerful incumbents in AI and reduce competition, thus stymieing innovation and concentrating AI's benefits."); *see also* Mark A. Lemley, *The Contradictions of Platform Regulation*, 1 J. FREE SPEECH L. 303, 335 (2021) ("[I]n the long run, regulatory choices that impose obligations on incumbents to do the things we want them to do as a matter of social policy are likely to entrench those incumbents, making it harder and more costly for someone to compete with them and eliminating the possibility of competing by offering a different set of policies"); LAWRENCE LESSIG, FREE CULTURE: HOW BIG MEDIA USES TECHNOLOGY AND THE LAW TO LOCK DOWN CULTURE AND CONTROL CREATIVITY 199 (2004) ("Overregulation stifles creativity. It smothers innovation. It gives dinosaurs a veto over the future. It wastes the extraordinary opportunity for a democratic creativity that digital technology enables.").

[216] *See* Zhang, *supra* note 12, at 120.

those potentially competing values. Therefore, policymakers should take agility, robustness, and proportionality into consideration and integrate these objectives into the policy framework design for AI and other fast-evolving technologies.

VII. CONCLUDING REMARKS

This research targets at worldwide policymakers who are struggling to frame advanced regulatory solutions to AI issues. The article helps policymakers navigate through the complex state of AI regulation and offers them an implementable three-layered approach to design a customized AI governance framework in a specific field for a specific jurisdiction. This three-layered framework maps various regulatory toolboxes and tools buttressed by well-selected studies and cases for policymakers to learn. We argue that the EU's model serves merely as one blueprint for global AI governance, instead of a one-size-fits-all standard that can be directly applied in other jurisdictions. This article offers a window for different jurisdictions' policymakers to better understand and evaluate those pioneers––including the EU––in AI regulation. Our methodology, in the same vein, can be used as a template to analyze other jurisdictions' legislation, facilitating global regulatory experimentation and learning. [217] Moreover, the prescriptions that we offer for each regulatory tool can help judges better understand each tool's role and placement within the broader regulatory landscape, enabling them to interpret and apply the rules in a more accurate way.

Theoretically, this research systemizes traditional regulation theories and applies those distilled lessons to a new context. Specifically, it revives certain overlooked rules and clarifies some misunderstandings about risk-based AI governance. First, different regulatory tools possess different functions and are subject to different realities; thus, the

[217] *See* Viktor Mayer-Schönberger & Urs Gasser, *A Realist Perspective on AI Regulation*, FOREIGN POL'Y (Sep. 24, 2024), https://foreignpolicy.com/2024/09/16/realist-perspective-ai-regulation/.

regulatory tools should be evaluated and implemented case by case. Generalizing about whether regulation is good or bad risks oversimplification. Second, policymakers and researchers should not ignore the critical possibility that the market mechanism has great inherent potential to mitigate AI and other emerging risks. And, market-based regulatory tools can coexist harmoniously with value-directed and procedural control rules. Good AI governance is not about selecting a single best regulatory tool, but about strategically integrating and harmonizing various tools to achieve an arguably best outcome.

This general policy guide will also benefit and benefit from further granular research. First, jurisdiction-specific research can enrich our understanding of how the three-layered framework is applied to a specific emerging industry, like robotics, as well as the impacts of the very regulatory arrangements. Second, policymakers may consider coordination and cooperation among regulatory institutions that are responsible for different layers, so that unitary regulatory objectives can be eventually achieved. Third, comparative research on diverse jurisdictions with similarities in some respects (*e.g.*, political regime, market conditions, and culture) but differences in the mounting of regulatory tools will provide insights into what constitutes a better regulatory mode. Lastly, international law researchers can draw on this three-layered framework to generate more detailed governance guidance for countries at different stages of AI development.

APPENDIX

Objectives	Layers		Toolboxes	Tools
Agility & Proportionality	**1st Layer** Market-Invigorating Strategies (Part II)	*Regular & Iterative Risk Assessment & Regulation Adjustments* (Part V)	Consumer Empowerment	Disclosure-to-Consumer (Transparency)
				AI Literacy
				Rights and Responsibilities (R&R)*
				Others
			Competition Adjustment	Competition Law/ Antitrust Law
				Open Sources Rules
				Non-Discrimination / Limiting Vertical Integration
				Break-Ups
				Interoperability
				Data Access
				Management of Labor Flow
				Regulatory Sandbox*
				Others
Robustness & Proportionality	**2nd Layer** Value-Directed Rules (Part III)		Command-and-Control (C&C)	Licensure
				Standard-Setting
				Regulatory Sandbox*
				Conduct-Based Duties / Risk Mitigation Duties / Rights Protection Duties
				Prohibition & Publishment
				Others
			Rights and Responsibilities (R&R)*	Life and Health Protection
				Non-Discrimination
				Right to Privacy
				Right to Likeness
				Right to Reputation
				Copyright
				Others
Robustness	**3rd Layer** Procedural Controls (Part IV)		Internal Management	Various
			Information Duty	Documentation
				Disclosure-to-Regulator
				Others