Regulatory
Studies Center
THE GEORGE WASHINGTON UNIVERSITY

ADAPTIVE REGULATION OF EMERGING TECHNOLOGIES, RISKS AND REACTIONS: FROM CLIMATE TO AI

Arden Rowell & Jonathan Wiener

SEPTEMBER 2026

© 2026 by author(s). All rights reserved.

The GW Regulatory Studies Center is an academic center in the George Washington University and its Trachtenberg School of Public Policy and Public Administration. This working paper is part of a series on Regulatory Approaches to Emerging Technologies supported by the John Templeton Foundation. Any views expressed in this working paper are those of the authors and not necessarily those of the John Templeton Foundation, the GW Regulatory Studies Center, the University of Illinois, or Duke University.

https://regulatorystudies.columbian.gwu.edu/

ABSTRACT

Emerging technologies present distinctive challenges for risk regulation, including (i) uncertainty in the face of low and emerging information, (ii) iterative dynamic change, (iii) the possibility of extreme tail risks and opportunities, and (iv) strategic or systemic reactions to regulatory intervention. This paper argues that these features jointly undermine static regulatory approaches and support adaptive regulation as the default framework for governing emerging technologies.

Adaptive regulatory strategies promote ongoing learning and policy updating, making them especially well-suited to technologies characterized by evolving information profiles. However, adaptive learning alone may be insufficient in the presence of extreme tail outcomes, where learning from experience may be impossible, and where the rareness of the outcome precludes iterative updating.

We therefore extend adaptive regulation in two directions. First, we propose explicit attention to extreme tail outcomes, emphasizing the need to identify whether catastrophic and beneficial tails are symmetrical or asymmetric. Second, we analyze "reactive" technologies—those that respond strategically or evolutionarily to regulation—and argue that effective governance must combine mitigation strategies that reduce risk probability with adaptation strategies that reduce the magnitude of harm. We discuss examples of emerging risks exhibiting some or all of these four characteristics, including climate change and AI/AGI, and we outline how the use of adaptive strategies—with ongoing updating, attention to extreme tail impacts, and combining mitigation (prevention) and adaptation (coping, resilience)—can help deal with emerging risks and reactions. In the case of climate change, slow progress on mitigation of greenhouse gas emissions, and rising climate impacts, have accentuated interest in climate adaptation (and sunlight reflection) to cope with climate risks. As to AGI, our framework suggests that a mitigation-only regulatory strategy—even if precautionary against its uncertain, dynamic and extreme tails, and even if adaptive over time—may still be vulnerable to AGI's reactions to such mitigation (e.g. circumvention or backlash), and thus a strategy combining both mitigation (prevention) and adaptation (coping or coexistence) may be optimal. Hence, we suggest a societal consequentialist reason (complementary to but distinct from reasons rooted in respect or rights for individuals) to develop adaptive regulation of AGI that avoids adverse reactivity and reduces existential risks to human survival.

AUTHORS

Arden Rowell is Associate Dean for Research, Mildred van Voorhis Professor of Law and Affiliate of the Beckman Institute for Advanced Science and Technology at the University of Illinois, and Jonathan Wiener is the Perkins Professor of Law, Professor of Public Policy and Environmental Policy, and Co-director of the Center on Risk, at Duke University, and a University Fellow at Resources for the Future (RFF).

ACKNOWLEDGEMENTS

This project/ publication was made possible through the support of Grant 63641 from the John Templeton Foundation. The opinions expressed in this publication are those of the authors and presented for academic discussion. They do not necessarily reflect the views of the John Templeton Foundation, the GW Regulatory Studies Center, the University of Illinois, or Duke University.

For helpful comments and discussion, the authors thank Susan Dudley, Sara Gerke, Chase Hamilton, Sarah Lawsky, Gary Marchant, JB Ruhl, Peter Salib, Bill Watson, Verity Winship, and participants in workshops at the George Washington University Regulatory Studies Center and the University of Illinois.

ABOUT THE GEORGE WASHINGTON UNIVERSITY REGULATORY STUDIES CENTER

Established in 2009, the GW Regulatory Studies Center is an academic center of the George Washington University and its Trachtenberg School of Public Policy and Public Administration. The Center's mission is to improve regulatory policy through research, education, and outreach. The Center is a leading source for applied scholarship in regulatory issues, and a training ground for anyone who wants to understand the effects of regulation and ensure that regulatory policies are designed in the public interest.

I. Introduction

Technologies—applications of knowledge to convert inputs into outputs, to solve problems, and to generate new opportunities—carry the potential for both benefits and risks to individuals and society. A technology can be considered a bundle or portfolio of opportunities and risks—and an emerging technology represents a new emerging bundle. Risk regulation in response to emerging technologies also represents a bundle or portfolio of opportunities (risk reduction) and risks (countervailing harms)—one can even consider regulation as a technology, a human construct to address the risks and opportunities of our other technologies.[1]

Ideally, the policy of risk regulation would be based upon complete information about the type, likelihood, magnitude, and distribution of both benefits and risks associated with a technology bundle. But while technologies are emerging, such information is limited. A classic pattern is that law lags technology: technologies emerge, then laws respond, and laws are always trying to keep pace with technological change.[2]

This paper begins in Part II by identifying features that are distinctive about risk regulation of emerging technologies. It argues that emerging technologies often exhibit at least four characteristics that affect the best approach to risk regulation: (1) initially low information (high uncertainty); (2) iterative dynamic change over time; (3) the possibility of extreme tail risks and/or opportunities; and (4) the possibility of reactions to regulatory policies. Each of these distinctive features poses challenges for risk regulation and can warrant special regulatory approaches.

Part III develops recommended regulatory responses. It begins by recommending that adaptive regulation be adopted as the default approach to regulating emerging technologies. Adaptive regulatory processes offer a calibrated approach for managing emerging technologies' initial low information and iterative dynamic change over time. These processes allow for incorporation of ongoing learning and policy updating, and thus are well-suited to adjusting to new information (e.g. via ex post or retrospective review) and integrating iterative dynamic change over time (e.g. via periodic review and ongoing sequential updating). We also make recommendations where emerging technologies present extreme tail risks and/or opportunities. In such cases, we recommend first specifying the tail(s), and then explore strategies for addressing these tails, for example through scenario-based planning as an aid for improving future imaginability before such tails become manifest. Finally, we label the category of "reactive" risks, arguing that where emerging technologies or their operators may react to regulation (such as by circumvention or

[1] See Jonathan B. Wiener, "The Regulation of Technology, and the Technology of Regulation," 26 *Technology in Society* 483-500 (2004).
[2] Gary E. Marchant et al., eds., *Innovative Governance Models for Emerging Technologies* (Northampton, MA: Edward Elgar, 2014); Gary E. Marchant, "The Growing Gap Between Emerging Technologies and the Law," in Gary E. Marchant, Braden R. Allenby, and Joseph R. Herkert eds., *The Growing Gap Between Emerging Technologies and Legal-Ethical Oversight: The Pacing Problem* 19 (Dordrecht, Netherlands: Springer, 2011).

backlash), optimal regulatory design should employ regulatory approaches to manage risk not only via mitigation (prevention) but also via adaptation (coping or coexistence).

Part IV concludes.

II. Distinctive Features of Emerging Technologies

a. Emerging Information Profiles and Dynamic Change

Risks are always uncertain to some degree because we can never fully know the future. Managing risk has been essential to human survival over millennia.[3] Risk policy works best when it can rely upon accurate information for foresight about the type, likelihood, magnitude, and distribution of risks and opportunities.

In the real world, perfect information about each of these factors is elusive (or illusory). Even for familiar risks from established technologies, such as automobile accidents, we face uncertainty about foreseeing risk factors such as driver behavior, distractions, vehicle performance, road conditions, and weather. New technologies may pose new risks and new uncertainties. Although often modeled as two separate categories of known vs. unknown risks, or risk with defined probability and uncertainty with undefined probability,[4] one can also model risks presented by new technologies as falling along a spectrum, where information profiles (and related uncertainty) range from low to high. A spectrum of less- to more-complete information profiles allows us to conceive of risks as temporally situated—that is, a construction of the information we have gathered at any particular point in time, and which may well change in future. For emerging technologies in particular—whose information profiles are especially prone to change over time—this is an important feature.

The creation of an information profile can be deliberate or accidental. Sometimes new information is revealed through surprise—unexpected discovery of better understanding of the technology and its risks—perhaps in the laboratory, or perhaps though an accident or disaster that shocks scientists and policy makers. And there can even be uncertainty about where each technology or risk sits on this spectrum; consider, for example, chlorofluorocarbons (CFCs), which were initially thought to be chemically inert (non-reactive) and thus a safe substitute for toxic compounds (such as ammonia) previously used in refrigerators and other applications; only decades later were CFCs recognized to be susceptible of rising into the stratosphere, there being broken apart by ultraviolet radiation from the Sun, and then contributing to the depletion of the protective stratospheric ozone layer.

[3] See Peter Bernstein, *Against the Gods: The Remarkable Story of Risk* (John Wiley, 1996).
[4] Cf. Cass R. Sunstein, "Knightian uncertainty in the regulatory context," 9 *Behavioural Public Policy* 614-629 (2025), doi:10.1017/bpp.2024.59 (evaluating distinctions between risk versus uncertainty, and approaches to decision-making with and without defined probabilities).

One key factor shaping the information profile is the time there has been to study and learn about a particular risk or technology. Emerging technologies are novel, and thus they can be expected to start with low information profiles. Uncertainty is often the basis for calls for ex ante precaution, but if all risks are uncertain and all risk regulation entails decision-making under uncertainty, then such calls for (stringent) precaution can be overbroad, especially if they would pose their own risks.[5] A better perspective is that uncertainty warrants inquiry: uncertainty is an opportunity for learning.[6]

Many emerging technologies will eventually develop past this early stage; they will be adopted and become established technologies with established operational data and actuarial risk information. In this sense, the low information profile is a characteristic of emerging technologies only so long as they remain in the emerging stage.

Other emerging technologies may never develop past their low information profile. One reason for this would be because they never come into use, having been curtailed by regulatory policy or market competition. Other emerging technologies may continue to be contested, with risk profiles that remain uncertain and unfolding.

The challenge is not only that emerging technologies can exhibit low information (high uncertainty). As noted, all technologies and all risks are uncertain to some degree, because the extreme end of the spectrum—perfect knowledge—is not achievable. It is only a first, and approximate, answer to treat uncertainty as an initial (ex ante) condition which can be addressed by policy precaution and then overcome by further research and experience toward subsequent (ex post) retrospective review. The challenge of emerging technologies goes beyond this two-stage ex ante / ex post framing, to recognize that things change—and keep changing.[7] Changes in scientific understanding can reduce uncertainty (potentially revealing lower or higher risk or opportunity than understood ex ante), but meanwhile changes in the technology, the environment, public policies, social patterns, public values, and other factors may continually change the risks and benefits of the technology. Such iterative dynamic change challenges the

[5] See Jonathan B. Wiener, "Precautionary Principle," chapter VI.13, pp.174-185, in volume 6, *Principles of Environmental Law* (Ludwig Kramer and Emanuela Orlando, eds.) of the *Encyclopedia of Environmental Law* (Michael Faure, ed.) (IUCN and Edward Elgar, 2018); Jonathan B. Wiener, "Precaution in a Multirisk World," in Dennis Paustenbach, ed., *Human and Ecological Risk Assessment: Theory and Practice* 1509-1531 (New York: John Wiley & Sons, 2002), reprinted in John S. Applegate, ed., *Environmental Risk* (London: Ashgate Publishing, 2004); Gregory E. Kaebnick et al., "Precaution and Governance of Emerging Technologies," 354 *Science* 710-711 (November 10, 2016).

[6] Yair Listokin, "Learning through Policy Variation," 108 *Yale L.J.* 480 (2008).

[7] See Gary Marchant et al., supra ; Lori S. Bennear and Jonathan B. Wiener, "Built to Learn: From Static to Adaptive Environmental Policy," in Daniel C. Esty, ed., *A Better Planet: Forty Big Ideas for a Sustainable Future* (Yale Univ. Press, 2019); JB Ruhl, "Adaptive Governance Law in Theory and Practice: Phantom or Phoenix?" 9 *J. Law & Innovation* 1-26 (2026).

institutions of risk regulation to develop sequential processes of ongoing learning and updating—planned adaptive regulation.[8]

b. Emerging Tails of Extreme Risk and Opportunity

Another quality of emerging technology that matters to risk regulation is that emerging technology may present extreme tails—whether via catastrophic risks,[9] amazing opportunities (such as to reduce risks),[10] or both.[11] Old tools designed to address existing circumstances and technologies may not always be sufficient to deal with these new types and tails of risk and opportunity.

Consider, on this front, the emergence of nuclear weapons technology and its global catastrophic risk—the risk of civilizational collapse or human extinction through large-scale nuclear war. This was a different type of risk than had been generated by earlier weapons technologies.[12] Managing it through the Cold War included a number of innovations in risk management, including deterrence strategies (which may also have prevented further conventional wars among great powers), information sharing methods such as a "Hotline" for real-time direct communication between nuclear powers, and a series of arms control and risk reduction treaties.[13]

Some technologies can generate new tails of opportunity and risk simultaneously.[14] Consider the development of new biomedical technologies, such as the invention of penicillin (antibiotics);[15]

[8] See Bennear and Wiener (2019), supra; Lawrence E. McCray, Kenneth A. Oye, and Arthur C. Petersen, "Planned Adaptation in Risk Regulation," 77 *Technological Forecasting & Social Change* 951–59 (2010); Marc A. Saner, "The Role of Adaptation in the Governance of Emerging Technologies," in Gary Marchant et al., eds., *Innovative Governance Models for Emerging Technologies* 106 (Northampton, MA: Edward Elgar, 2014); Thibault Schrepel, "Adaptive Regulation," — *European J. Risk Regulation* — (2026), doi:10.1017/err.2026.10087 .

[9] See Cass Sunstein, *Worst Case Scenarios* (2009); Nick Bostrom, "Existential Risk Prevention as Global Priority," 4 *Global Policy* 15-31 (2013); Toby Ord, *The Precipice* (2020); Elisabeth Paté-Cornell, "On Black Swans and Perfect Storms: Risk Analysis and Management When Statistics Are Not Enough," 32 *Risk Analysis* 1823-33 (2012). On the obstacles to mitigating extreme catastrophic risks, see Jonathan B. Wiener, "The Tragedy of the Uncommons," 7 *Global Policy* 67-80 (2016).

[10] See Nick Bostrom, "Astronomical Waste: The Opportunity Cost of Delayed Technological Development," 15 *Utilitas* 308 (2003); Arden Rowell, "Regulating Best-Case Scenarios," 50 *Environmental Law* 1105–1172 (2020).

[11] *Id.*

[12] See Martin Sherwin, *A World Destroyed: Hiroshima and Its Legacies* (1975).

[13] See Lawrence Freedman, *The Evolution of Nuclear Strategy* (4th ed. 2019).

[14] See Rowell (2020), *supra.*

[15] See Robert Bud, *Penicillin: Triumph and Tragedy* (2009).

genetic engineering of plants and animals;[16] or geoengineering.[17] The transformative impacts of such technologies are extraordinary—potentially saving many lives—even as their development may generate new risks such as antibiotic-resistant pathogens and adverse ecological impacts. Or consider the opportunities and risks presented by emerging technologies for space exploration, which offer the potential of expanding scientific knowledge, pursuing substantial resource extraction (e.g. through asteroid mining), and supporting long-term habitation beyond Earth—while also introducing risks such as forward-and backward-contamination (interplanetary movement of microbes), contact with dangerous extraterrestrial life, and the militarization or corporatization of space.[18] This dynamic illustrates a broader pattern: even where emerging technologies generate net benefits that dramatically improve human welfare, they may simultaneously alter underlying biological, environmental, or social systems in ways that generate novel (and sometimes compounding) risks.[19]

c. Emerging Interactions with Regulation: Reactivity

The novelty of emerging technology also matters to risk regulation if it can interact with the regulation, or what we call "reactivity." Reactive new technologies and actors can respond iteratively to regulatory policies, such as through circumvention or backlash. This feature adds a fourth challenge, beyond uncertainty (low information), dynamic change, and extreme tails. Interactivity between an emerging technology and regulatory tools can be positive and/or negative.

In some cases, the development of new technologies opens new opportunities for better regulation, such as by reducing cost, increasing effectiveness, or serving some other value (such as increasing democratic accountability). Consider the emergence of remote sensing technologies

---

[16] See National Academies of Sciences Engineering and Medicine, *Genetically Engineered Crops: Experiences and Prospects* (2016); Gary Marchant, Braden R. Allenby, and Joseph R. Herkert (2004), *supra.* Even if existing widely used methods of genetic modification or gene editing are now thought *not* to pose extreme catastrophic risks, and may reduce other risks to health and environment, see Sunstein, "Knightian uncertainty," *supra* n.4 at 621 ("the examples of genetically-modified crops and nuclear power look badly out of date"), some leading scientists are nonetheless concerned that new methods possibly enabling the synthesis of opposite-chiral ("mirror") organisms may pose extreme catastrophic risks, see Katarzyna P. Adamala et al., "Confronting Risks of Mirror Life," 386 *Science* 1351-1353 (12 Dec. 2024), DOI: 10.1126/science.ads9158 .

[17] For early appraisals of this emerging technology and its regulation, see Scott Barrett, "The Incredible Economics of Geoengineering," 39 *Environmental and Resource Economics* 45-54 (2008); David G. Victor, "On the Regulation of Geoengineering," 24 *Oxford Review of Economic Policy* 322-336 (2008); David G. Victor et al., "The Geoengineering Option: A Last Resort Against Global Warming?" 88 *Foreign Aff.* 64 (2009); Daniel Bodansky, "The Who, What, and Wherefore of Geoengineering Governance," 121 *Climatic Change* 539 (2013); Albert Lin, "The Missing Pieces of Geoengineering Research Governance," 100 *Minn. L. Rev.* 2509 (2016). For recent analyses of the benefits, risks and regulation of solar geoengineering, see Tyler Felgenhauer et al., "Practical Paths to Risk-Risk Analysis of Solar Radiation Modification," 5(1) *Oxford Open Climate Change* kgaf012 (2025), https://doi.org/10.1093/oxfclm/kgaf012 ; Jonathan B. Wiener et al., "Multi-Risk Governance of Solar Radiation Modification," 16 *European J. Risk Regulation* 1246-1262 (2025), https://doi.org/10.1017/err.2025.10018 .

[18] See Jonathan B. Wiener and Chase Hamilton, "Interplanetary Risk Regulation," 26 *Chicago J. Int'l Law* 73-106 (2025).

[19] See Rowell (2020), *supra.*

for environmental monitoring, which reduced the cost of monitoring and enabled better frequency of measurements, making performance-based regulatory tools less costly and more effective to adopt and enforce.[20] Or consider the emergence of open data portals and their impacts on public accessibility and auditing of government action.[21]

Other interactions between emerging technologies and regulatory tools can be negative, insofar as they make regulation more costly, less effective, or pose countervailing risk tradeoffs. Consider, for example, how encryption technologies can undermine traditional monitoring approaches and make it harder to enforce rules on improper transactions, national security, or illicit online content.[22] Or how ride-sharing apps reorganized taxi and transportation services, undermining traditional economic regulation of taxis and disrupting the incumbent taxi industry.[23]

The impact of the interaction between a new technology and the cost or effectiveness of a regulatory tool may be so great that it changes optimal policy design. For instance, regulatory tools that developed for prior technologies can become obsolete.[24] This can be observed when technology-based standards impair innovation and make improved performance costly. For example, existing technology-based nuclear safety regulations were designed for light-water reactors (LWRs); these prescriptions are often obsolete or inapplicable when applied to emerging technology in advanced reactors.[25] Safety measures designed to prevent meltdown of solid fuel (such as happened at Fukushima, an LWR), for example, are inapplicable to molten salt reactors, whose fuel is already molten.[26]

The emergence of a new technology can also make some regulatory tools newly attractive, as where the technology lowers the cost or increases the effectiveness of tools that were not feasible or sensible to use prior to development of that technology. Consider, for example, the rollout of smart electricity meters, which have made dynamic energy pricing feasible,[27] or the application of blockchain and advanced "MRV" (monitoring, reporting, and verification) systems to improve

[20] See Cary Coglianese & David Lazer, *Management-Based Regulation: Prescribing Private Management to Achieve Public Goals*, 37 Law & Soc'y Rev. 691 (2003).
[21] See Archon Fung, Mary Graham & David Weil, *Full Disclosure: The Perils and Promise of Transparency* 3-28 (2007).
[22] See Peter Swire & Kenesa Ahmad, *Encryption and Globalization*, 13 Colum. Sci. & Tech. L. Rev. 416 (2012).
[23] See Peter Molk and Arden Rowell, *Reregulation and the Regulatory Timeline*, 101 Iowa L. Rev. 1497 (2016).
[24] See Richard B. Stewart, *Regulation, Innovation, and Administrative Law: A Conceptual Framework*, 69 Calif. L. Rev. 1256 (1981)
[25] See George Joslin et al., *Managing Uncertainty in Efficiency, Safety, and Independence for the Licensing and Regulation of Advanced Nuclear Reactors*, _ Nuclear Science and Engineering (2026).
[26] See Badawy M. Elsheikh, *Safety assessment of molten salt reactors in comparison with light water reactors*, 6 J. of Radiation Research and Applied Sciences 63 (2013).
[27] See Paul L. Joskow & Catherine D. Wolfram, *Dynamic Pricing of Electricity*, 102 American Economic Review 381 (2012).

emissions tracking, increasing the viability of cap-and-trade / emissions trading systems (ETS) to reduce greenhouse gas emissions.[28]

Still other interactions blend positive and negative effects. Consider the way in which machine learning enables individualized pricing practices, allowing market pricing to shift according to individual willingness to pay. This may both improve and worsen the policy impacts of market-based regulation—by making exchanges more efficient while simultaneously creating a tool that can be used for price discrimination.

III. Risk Management of Emerging Technologies

Regulation of emerging technologies should account for the distinctive challenges they present. Here we recommend risk management strategies to address the four features discussed above. First, we recommend that adaptive regulatory strategies be adopted as the default for emerging technologies, to deal with low information (uncertainty) and dynamic change. Then we recommend approaches to address the extreme tail risks and opportunities that emerging technologies may present. Finally, we recommend approaches to address reactivity, including reactivity by AI, where emerging technologies interact with and may react to regulation.

a. Adaptive Defaults for Low Information and Dynamic Change

Conventional regulation often struggles or fails to deal with emerging technologies because it takes a static approach to a dynamic phenomenon. A major regulation is often a one-time decision, or requires a long time to revise, while the technologies and the world continue to change rapidly—leaving the initial policy obsolete, or an impediment to innovation, while the actual realized benefits, costs and risks diverge from the ex ante projections.[29] Consider a spectrum from perfectly static regulation (once adopted, it never changes, no matter how much new information) to perfectly adaptive regulation (it updates continuously and instantaneously in response to every iota of new information): neither end of this spectrum would be optimal. Static policy "mismatches" with a changing world can lead to policy failures[30]—excessive costs, insufficient benefits, public frustration and loss of trust in institutions, and accidents or disasters (spurring "policy shock" and then un-planned hasty policy revision[31]). But a perfectly adaptive policy would incur excessive costs of monitoring, analysis, and policy instability. Adaptive regulation is thus needed to incorporate ongoing learning and updating into dynamic policy for dynamic technologies and conditions.

---

[28] See World Bank, *Blockchain and Emerging Digital Technologies for Enhancing Post-2020 Climate Markets* (2018).
[29] See Marchant (2014), supra; Bennear and Wiener (2019), supra.
[30] See Bennear and Wiener (2019), supra.
[31] See Edward J. Balleisen, Lori S. Bennear, Kimberly D. Krawiec, and Jonathan B. Wiener, eds., *Policy Shock: Recalibrating Risk and Regulation after Oil Spills, Nuclear Accidents, and Financial Crises* (Cambridge: Cambridge University Press, 2017).

Adaptive regulatory strategies—which allow for learning and modification of policy over time via adjustments informed by data collection and analysis—are a strong fit where net benefits profiles are likely to change considerably (hence a high value of learning to improve policy) and where the costs of adaptive processes (such as analysis, deliberation, and policy instability) are kept reasonable.[32] While it is possible that some emerging technologies have unchanging net benefits profiles and sufficient information available to set policy options ex ante, such cases are likely to be rare for important emerging technologies. As a result, adaptive regulatory strategies offer the most appropriate default for emerging technologies.

To deal with low information (high uncertainty), ex ante regulation can and should use the best methods of foresight available, including forecasting, risk assessment, regulatory impact assessment, scenario-based planning, and expert elicitation.[33] Regulatory impact assessment (at every stage) needs to take a broad scope, examining not only intended impacts but also unintended or ancillary impacts, such as countervailing risks and co-benefits.[34] Some of these ancillary impacts may be well foreseeable at the outset of an emerging technology and initial regulation, but others may emerge as the technology is tested and used and the information profile about its full effects rises from low to modest or high.

A first step toward adaptive regulation is to plan ex ante for (and follow through with) at least one subsequent round of ex post evaluation (retrospective review).[35] Such review will require planning ahead (ideally in the initial statute or agency rule) for monitoring, data collection, analysis of the policy and its impacts (intended and unintended/ancillary) compared to counterfactual scenarios of what would have happened otherwise, and a process to decide whether and how to update the policy in light of this richer information profile (including a new exercise in foresight about the next policy options). Retrospective review is often requested by presidents but not often undertaken by agencies; that may be because agencies lack the resources and the motivation to review their past rules. Equipping and pressing agencies to review each past rule, one at a time, can yield some improvements in each rule (such as lower costs, greater

[32] See Bennear and Wiener (2019), supra.
[33] See Jennifer Kuzma, et al., "An Integrated Approach to Oversight Assessment for Emerging Technologies," 28 *Risk Analysis* 1197-1219 (2008); Susan E. Dudley and Brian F. Mannix, "Improving Regulatory Benefit-Cost Analysis," *The Journal of Law and Politics* (2018). Public participation in ex ante regulatory analysis of emerging technologies can also provide important inputs to experts' evaluations and ongoing policy learning, including by supplying knowledge that corrects or improves experts' models and analyses, and by communicating public values through participatory mechanisms (complementary to the public values that are incorporated in policy analyses such as benefits valuations). See Albert Lin, "Technology Assessment 2.0: Revamping Our Approach to Emerging Technologies," 76 *Brooklyn L. Rev.* 1309 (2011); Jonathan B. Wiener, "Risk in the Republic," 8 *Duke Envt'l Law & Policy Forum* 1 (1997).
[34] See John D. Graham and Jonathan B. Wiener, *Risk vs. Risk* (Cambridge: Harvard Univ. Press, 1995); Jonathan B. Wiener, "Learning to Manage the Multi-Risk World," 40 *Risk Analysis* 2137-2143 (2020).
[35] See Maureen Cropper, Arthur Fraas and Richard Morgenstern. "Looking Backward to Move Regulations Forward," 355 *Science* 1375-1376 (2017); Susan Dudley and Sally Katzen, "Crossing the Aisle to Streamline Regulation," *Wall Street Journal*, May 13, 2019, at https://www.wsj.com/articles/crossing-the-aisle-to-streamline-regulation-11557788679 .

benefits, and avoiding adverse side effects). A broader approach may overcome agency inhibitions and yield deeper insights by charging an institution to review multiple rules across multiple agencies (such as a multi-agency group, an executive oversight body, or a new independent commission), thereby enabling comparisons across variations in regulatory approaches and variations in impact assessment methodologies, offering more comprehensive regulatory learning about optimal policy design and better methods of policy foresight.[36]

To deal with dynamic change, risk regulation can and should build mechanisms for ongoing policy learning over time via planned adaptive regulation.[37] This involves iterative sequential reviews, with iterative monitoring, data analysis, and updating. Mechanisms for such ongoing policy learning can include periodic review at designated intervals such as every X years (with the optimal time interval depending on the rate of change, the value of learning and policy updating, and the costs of revision).[38] Periodic review is often called for by federal statutes, yet incompletely implemented by agencies, perhaps because of the reasons noted above, including lack of resources and of motivation to review past rules; successful processes of periodic review can involve dedicated staff assigned to ongoing evaluation of new information.[39] Adaptive revisions can be discretionary (based on evaluation and judgment about best policy adjustments), and/or automated (based on pre-designated indicator variables such that the policy adjusts when threshold levels are reached). Discretionary adaptive regulation is more information-rich (considering more variables including intended and unintended effects) but requires more time and personnel to evaluate and decide, whereas automated adaptive regulation is faster and administratively cheaper but is more information-thin (limited to pre-designated indicator variables).[40]

Dynamic planned adaptive regulation shifts risk regulation from a one-time answer to a continuous process of inquiry. It shifts the intensity of regulatory analysis (and of judicial review of agency rulemaking) from up-front ex ante assessment to ongoing sequential reappraisals.[41] It depends on information about actual outcomes, entailing greater investments in monitoring and data analysis. Investment in such efforts at ongoing impact assessment and repeated periodic review[42] will be worthwhile in some but not all cases—especially where the emerging

---

[36] See Lori S. Bennear and Jonathan B. Wiener, "Institutional Roles and Goals for Retrospective Regulatory Analysis," 12 *Journal of Benefit-Cost Analysis* 466-493 (2021).
[37] See Bennear and Wiener (2019), supra; Rowell and Molk (2016), supra; McCray, Oye and Petersen, supra.
[38] See Lori S. Bennear and Jonathan B. Wiener, "Periodic Review of Agency Regulation," report for the Administrative Conference of the United States (ACUS) (June 7, 2021), at https://www.acus.gov/report/periodic-retrospective-review-report-final.
[39] Id.
[40] See Bennear and Wiener (2019), supra.
[41] See Robin Craig and J.B. Ruhl, "Designing Administrative Law for Adaptive Management," 67 *Vanderbilt Law Review* 1-87 (2014); Wendy Wagner et al., "Dynamic Rulemaking," 92 *NYU Law Review* 182-266 (2017); Daniel A. Farber, "Environmental Protection as a Learning Experience," 27 *Loyola L.A. Law Review* 791–807 (1994); Justin R. Pidot, "Governance and Uncertainty," 37 *Cardozo Law Review* 112-184 (2015).
[42] See Daniel Lima Ribeiro, "Adaptive Regulatory Impact Assessment: Beyond the Foresight-Hindsight Divide," SJD Dissertation, Duke Law School (May 2018).

technology and other factors are likely to continue changing (hence high benefits of ongoing learning and updating), and where the costs of analysis and policy revision are reasonable.[43]

b. Addressing Extreme Tail Risks and Opportunities

Because emerging technologies introduce new and dynamic features to complex systems, they can solve old problems and present new opportunities (disrupting past systems in waves of "creative destruction"[44]), while also presenting new countervailing risks. Risk regulation faces distinctive pressures when attempting to address extreme novel benefits and risks. As noted above, regulatory impact assessment needs to take a broad scope that examines not only intended impacts (benefits) but also unintended ancillary impacts (co-benefits and countervailing harms)[45].

In addition, beyond impacts that are ordinary or easily foreseeable, emerging technologies may introduce potentially extreme impacts that are low-probability but high-consequence. These extreme impacts may lie at the far "tails" of the distribution of expected impacts.[46] Even if unlikely, their enormous consequences can make them important to policy decisions.[47] The extreme tail impacts of emerging technologies can be negative (global catastrophic or existential risks that could end life, or end human life, or destroy civilization), or positive (solving previously insoluble problems, preventing catastrophic/existential risks, generating seemingly miraculous benefits), or sometimes both.[48]

In theory, risk regulation can be designed to try to address such extreme tail impacts.[49] In practice, risk regulation often struggles to deal with extreme tails.[50] Public attention and policy maker decisions may neglect extreme impacts for several reasons, including that these effects are so rare that they are not salient in our experience (the "unavailability" heuristic[51]); are so large in magnitude that people feel overwhelmed and ineffective and hence underinvest in prevention ("mass numbing" or "compassion fade"[52]); or are so transformative or devastating that they are

[43] See Bennear and Wiener (2019), supra (identifying benefits and costs of adaptive regulation); Ruhl (2026), supra (praising adaptive governance but worrying that its costs lead governments to adopt it too rarely); Gary E. Marchant, "Governance of Emerging Technologies as a Wicked Problem," 73 *Vanderbilt Law Review* 1862 (2020).
[44] See Joseph A. Schumpeter, *Capitalism, Socialism, and Democracy* (Harper, 3rd ed. 1950).
[45] See Graham & Wiener (1995), supra; Wiener (2020), supra.
[46] See Martin Weitzman, *On Modeling and Interpreting the Economics of Catastrophic Climate Change*, 91 Rev. Econ. & Stat. 1 (2009); Robert S. Pindyck, *Uncertain Outcomes and Climate Change* Policy, 7 J. Envtl. & Econ. Mgmt. 152 (2011); Paté-Cornell, supra.
[47] See Weitzman (2009), supra; Pindyck (2011), supra; Bostrom (2013), supra; Wiener (2016), supra; Cass R. Sunstein, "Irreversible and Catastrophic," 91 *Cornell L. Rev.* 841 (2006); Rowell (2020), supra.
[48] See Rowell (2020), *supra*.
[49] See e.g. Jason Matheny, "Reducing the Risk of Human Extinction," 27 *Risk Analysis* 1335-1344 (2007); Daniel Farber, "Complexity Theory and Environmental Uncertainty," 37 *U.C. Davis L. Rev.* 145 (2003).
[50] See Jonathan B. Wiener, "The Tragedy of the Uncommons," 7 *Global Policy* 67-80 (2016).
[51] See Cass R. Sunstein, "Irreversible and Catastrophic," *supra* n.47, at 884 (2007) (mentioning "unavailability bias"); Wiener, "Tragedy of the Uncommons," *supra* (on "unavailability heuristic" and catastrophic risk).
[52] See Paul Slovic, "'If I look at the mass I will never act': Psychic numbing and genocide," 2 *Judgment and Decision Making* 79–95 (2007). at https://doi.org/10.1017/S1930297500000061 ; Paul Slovic and Daniel Vastfjall,

under-deterred by ex post remedies such as civil tort liability, criminal law sanctions, or insurance, which would be disabled by the manifestation of the transformation (e.g., the catastrophe would destroy these institutions, or the transformation would obviate them) and thus not credible ex ante.[53] Moreover, adaptive processes of ongoing regulatory learning may not be sufficient for extreme tails (such as downside risks of global catastrophe), for which learning from experience would be precluded if the risk manifests (e.g. ending all life, or at least disabling the institutions for subsequent learning and policy updating).[54]

One important puzzle in regard to the management of tails is whether both types of tails—catastrophes and miracles—should be treated symmetrically. On the one hand, neglecting one side of the equation can introduce policy distortion, institutionalizing loss aversion or benefits seeking.[55] And similarities between extreme upsides and downsides—including the size of impact; the role of risk and uncertainty; intertemporal distributions; and the trigger of cognitive phenomena[56]—may further militate for symmetric treatment. On the other hand, upsides and downsides also have important differences: catastrophes are bad while miracles are good; catastrophes are likely to be viewed as losses from the status quo, while miracles are likely to be viewed as gains; catastrophes make people poorer and sicker (or even extinct) while miracles make them wealthier and healthier (or even immortal); and catastrophes trigger fear and pessimism, while miracles implicate hope and optimism.[57] These differences may lead to differential treatment to upsides and downsides—though they do not necessarily all justify that different treatment.

Unfortunately, evidence is scant (and in many cases may be unattainable) on whether the public and policy makers are in fact more inattentive (i.e. neglectful of) one tail or the other, or whether they are equally inattentive to both extreme downside catastrophes and extreme upside opportunities. And the above factors may push in multiple directions simultaneously. Both extreme tails seem prone to neglect due to unavailability. Neglect of extreme downside catastrophes may be exacerbated by mass numbing, and underdeterrence (e.g. from failed institutions), as just noted, while other factors such as optimism bias and vain hope may lead people to inflate the likelihood of miraculous salvation such as by deus ex machina. On the other hand, neglect of extreme miraculous opportunities may be worse, if miracles are perceived as "gains" from the status quo, while catastrophes are perceived as "losses." If so, the behavioral phenomenon of "loss aversion" may be triggered for downside catastrophic risks (but not for upside opportunities), leading catastrophes to be systematically (and subconsciously) more

"The More Who Die, the Less We Care: Psychic Numbing and Genocide," in Susanne Kaul & David Kim, eds., *Imagining Human Rights* 55-68 (De Gruyter, 2015).

[53] See Wiener (2016), supra (discussing the three factors "unavailability," "numbing," and "underdeterrence" in fostering neglect of extreme catastrophic risks).

[54] Id.

[55] Rowell (2020), supra at 1168-69.

[56] Id. at 1146-1157.

[57] Id. at 1157-1164.

salient (if at all), more cognitively "available," more probable, and thus weighted more heavily than miracles—even when the magnitude of both is comparable.[58]

Under uncertainty on whether neglect of one type of tail (upside or downside) is more common generally, however, one useful policy intervention is to ensure that individual analyses at least consider both types. Doing so may help to avoid inadvertently introducing a policy skew (whether loss-averse or benefits-seeking) into analyses. Consider, for example, that people thinking or joking about extreme tail catastrophic/existential risks sometimes refer to "p(doom)"—the probability of human extinction.[59] Thus far apparently no counterpart upside—which we might style "p(miracle)" or "p(nirvana)"—has entered common parlance, as far as we've seen. A conversation that considers only doom, without countervailing miracles, is not just depressing; it runs the risk of inadvertently neglecting key tradeoffs and counterweights, and of unreflectively institutionalizing loss aversion. Insofar as inadvertent neglect of this type enters more formal policy analysis, it should be corrected, just as neglect of extreme existential catastrophes warrants correcting.

Where the policy analysis concerns emerging technology, it should further anticipate and manage both tails using adaptive methods that acknowledge and affirmatively manage limited information profiles and their potential change over time. In managing these profiles, it is important to affirmatively acknowledge that information may be limited about upside as well as downside tails—that is, that uncertainties may be positive, negative, or both. In managing interacting or simultaneous upside and downside tails, moreover, it is important for policymakers to consider the potential for inadvertent analytical bias noted above. It is common to propose precautionary measures for managing the catastrophic risks of emerging technologies.[60] Where such precautionary measures are adopted without consideration of the "other" tail—of the possibility of extreme upsides as well as extreme downsides—they risk introducing an inadvertent analytical skew into risk regulation. This failure to recognize both upside and downside tails is especially problematic where the tails interact with one another—as might be

[58] *Id.*; Daniel Kahneman & Amos Tversky, *Prospect Theory: An Analysis of Decision Under Risk*, 47 Econometrica 263, 263 (1979); Amos Tversky & Daniel Kahneman, Loss Aversion in Riskless Choice: A Reference-Dependent Model, 106 Q. J. ECON. 1039, 1039 (1991); Daniel Kahneman et al., Anomalies: The Endowment Effect, Loss Aversion, and Status Quo Bias, 5 J. ECON. PERSP (1991); William Samuelson & Richard Zeckhauser, Status Quo Bias in Decision Making, 1 J. RISK & UNCERTAINTY 7, 8 (1988).

[59] See Andrew Marantz, "Among the A.I. Doomsayers," *The New Yorker*, March 11, 2024 (published in the print edition of the March 18, 2024 issue, with the headline "O.K., Doomer."), at https://www.newyorker.com/magazine/2024/03/18/among-the-ai-doomsayers (reporting the "p(doom)" estimates of several AI industry experts); Kevin Roose, "AI Poses 'Risk of Extinction,' Industry Leaders Warn," *NY Times*, May 30, 2023, at https://www.nytimes.com/2023/05/30/technology/ai-threat-warning.html?smid=nytcore-ios-share ; David Streitfeld, "People Loved the Dot-Com Boom. The AI Boom, Not so Much," *NY Times*, Feb. 21, 2026, at https://www.nytimes.com/2026/02/21/technology/ai-boom-backlash.html (reporting on public fears of AI vs. tech industry experts' optimism for AI, but also that tech industry experts often say that AI will make humans obsolete). For more examples of tech insiders' estimates of p(doom), see https://en.wikipedia.org/wiki/P(doom)

[61] See Wiener and Hamilton (2025), supra.

the case with solar geoengineering, genetic modification, space exploration, artificial intelligence, and other emerging technologies posing global impacts.

Although the imaginability of extreme tails remains challenging, the high stakes of such tails suggests the need for investment in ex ante foresight, to be deployed before tails come to fruition (or fail to do so). Such foresight requires imagination. We do not typically experience extreme tail events; such events could destroy or reforge all that is familiar in our lives and institutions, and could either preclude further learning or trigger new ways of learning. To envision hypothetical future extremes before they occur, we might engage in structured scenario-based planning, and try to assess the likelihood and consequences of each scenario. In doing so, however, we should be attentive to the possibility of both upside and downside tails.

c. Managing Interactivity: Regulating Reactivity

A distinctive challenge of some technologies and risks is the potential for a kind of interactivity we call "reactivity." For some technologies, the sources of the technology and the risks being regulated may react to the regulation, whether intentionally strategically or via evolutionary processes, in ways that circumvent or complicate or thwart the regulatory policy, potentially yielding less effective or even perverse outcomes.

Consider several examples that pose uncertainty, dynamic change, and catastrophic risk, but vary in their extent of reactive feedbacks to regulation:

- Asteroids may pose emerging risks with initially low information (uncertainty), changing with improved detection, and potentially catastrophic impacts, yet they are not reactive to mitigation policies (not intentionally, although some deflection efforts may generate fragments or adjusted trajectories that perversely increase risk).[61]
- Microorganisms may pose emerging risks with initially low information (uncertainty) and potentially catastrophic impacts, plus dynamic migration and evolution over space and time, and highly reactive responses to mitigation efforts. They may evolve to overcome medical and agricultural prevention strategies, such as through selection pressures shaping the evolution of antibiotic-resistant or pesticide-resistant traits.
- Military adversaries and terrorist groups possessing weapons of mass destruction (WMD) may exhibit all four characteristics, including intentionally strategic responses to mitigation measures such as counterterrorism policies, e.g. by striking earlier or at alternative/softer targets.
- Chlorofluorocarbons (CFCs) illustrate the difficulty of early identification of some emerging technology risks: they were initially thought to be chemically inert and thus a superior replacement for toxic substances (e.g. ammonia) in refrigeration and heating/cooling, but new discoveries five decades later showed that this initial risk

[61] See Wiener and Hamilton (2025), supra.

assessment was based on limited information, because CFCs could rise in the atmosphere and be broken apart by solar ultraviolet (UV) radiation, thus freeing chlorine ions to deplete the protective stratospheric ozone layer, posing a global catastrophic risk. Subsequent adaptive regulations phased out CFCs, inviting industry to introduce a series of substitutes (e.g. HCFCs, HFCs, and beyond) posing a cascade of new risk tradeoffs (including global warming), and further adaptive regulatory policies.

- Climate change exhibits all four characteristics: uncertainty, dynamic change, extreme tail risks, and reactive feedback effects (such as sources reacting to evade regulatory limits on GHG emissions (e.g. "carbon leakage"), and feedbacks in the climate system itself). Some of the relevant technologies are old (e.g. burning fossil fuels) rather than newly emerging, but the substitutes for fossil energy can be emerging technologies (e.g. nuclear, renewables) with different risk profiles. Sunlight reflection methods (SRM) are emerging technologies that may counteract climate change but may also pose significant new countervailing environmental and geopolitical risks.[62]
- Artificial intelligence (AI)—such as in autonomous vehicles, autonomous weapons, autonomous robots on Earth or in outer space, automated government policy making, and other applications or decision settings—exhibits all four characteristics, arguably more than the other examples just noted. Notably, AI presents plausible scenarios generating both upside and downside tail risk. Furthermore, and distinctively, AI may react to prevention measures by evading (circumvention) or even by retaliating (backlash)—especially as AI evolves from task-specific skills to become artificial general intelligence (AGI) that can learn many tasks and even teach itself new skills (recursive self-improvement).[63]

For emerging technologies that generate reactive feedback effects, policymakers will need more than prevention/mitigation measures: they will also need adaptation strategies, to cope or coexist with strategic risk sources. Adaptive regulation needs to incorporate adaptation—ameliorating the magnitude of harm through resilience—as well as mitigation—reducing the probability of the scenario. Notably, where emerging technology risks significantly involve reactive feedback effects, then strategies to ameliorate the magnitude of harm should be added, and should accommodate upside as well as downside scenarios.

The last few years have seen significant advances in artificial intelligence (AI) technologies, and in attention to those technologies from the public and policymakers. Thus far, that attention has largely focused on AI systems that are designed for specific, well-defined tasks, without possessing the ability to generalize knowledge or adapt to diverse domains. The AI technologies that have been developed thus far, such as Large Language Models (LLMs), fall into this

[62] See Tyler Felgenhauer et al., "Practical Paths to Risk-Risk Analysis of Solar Radiation Modification," 5(1) *Oxford Open Climate Change* kgaf012 (2025); Jonathan B. Wiener et al., "Multi-Risk Governance of Solar Radiation Modification," 16 *European J. Risk Regulation* 1246-1262 (2025).
[63] See further discussion and citations below.

category. AI of this type—which may reach and even exceed a human level of performance, but only within a narrow area of specialization—is sometimes called Artificial Narrow Intelligence (ANI), to distinguish it from what is often seen as the next step in artificial technology: Artificial General Intelligence (AGI). AGI technologies are posited to differ from ANI in their ability to generalize—an ability to reach or surpass human-level of functioning across multiple domains, to learn new skills, to make decisions and develop strategies, to act autonomously, and to engage in general-purpose problem-solving.[64]

Although the social impact of ANI technologies has increased substantially in recent years, and much policy attention has been aimed at the near-term risks of ANI (e.g., labor market displacement, bias and discrimination, and misinformation),[65] the development of AGI—which some technologists believe could be coming quite soon[66]—poses substantially greater transformational potential. One important question relates to AGI's potential for existential risk—the risk of human extinction.[67] The extreme tail downside risk of AGI is that it might engage in conflict with humans (for instance, via competition for resources such as energy, or if AGI sees humans as an obstacle, or if humans seek to suppress AGI), or might seek to enslave or make pets of humanity, or might be weaponized into competing AGIs that engage in conflict with one another, with widespread human casualties or extinction. Such scenarios are sometimes treated as science fiction, but experts see them as possible enough that they justify policy concern: for example, the Center for AI Safety gathered numerous signatures for its "Statement on AI Risk," which warned that "[m]itigating the risk of extinction from AI should be a global priority alongside other societal-scale risks such as pandemics and nuclear war."[68]

---

[64] See Lisa Lacy, "There's AI, and then there's AGI," CNET (17 February 2024), at https://www.cnet.com/tech/theres-ai-and-then-theres-agi-what-you-need-to-know-to-tell-the-difference/ .

[65] See Charles I. Jones, "AI and Our Economic Future," NBER Working Paper 34779 (Jan. 2026), at https://www.nber.org/papers/w34779 . Kevin A. Bryan, "The Economic Impacts of Artificial Intelligence: A Multidisciplinary, Multi-book Review," 64 *Am. Econ. Rev*. 281-300 (Feb. 2026) (observing that economists are studying AI's economic impacts in the near term, but are giving "little guidance" on "transformative scenarios" such as "existential risk"; then concluding with "open questions" including "How can we, as with climate change, coordinate internationally on AI safety concerns? Game theorists have a role to play here."). Two centuries ago, Mary Wollstonecraft Shelley wrote the classic cautionary tale of human technology creating artificial human life, in *Frankenstein, or The Modern Prometheus* (1816); her preface to the 1831 edition warned of the dangers of what we now call AI.

[66] See Stanford HAI, *The 2026 AI Index Report* (April 2026), at https://hai.stanford.edu/ai-index/2026-ai-index-report ; Yohan J. John, "What Happens When AI Isn't Trained, But Evolves?" Kempner Institute, Harvard University (March 2, 2026), at https://kempnerinstitute.harvard.edu/news/what-happens-when-ai-isnt-trained-but-evolves/ (reporting on Joseph Bejjani et al., "The Emergence of Complex Behavior in Large-Scale Ecological Environments," arXiv (revised Dec. 2025), https://doi.org/10.48550/arXiv.2510.18221 , at https://arxiv.org/abs/2510.18221 .

[67] See Kevin Roose (2023), supra; Nick Bostrom, *Superintelligence* (Oxford Univ. Press, 2014); Matthew Tokson and Yonathan Arbel, "Artificial Intelligence and Existential Risk," – *Connecticut L. Rev.* – (draft March 26, 2026), at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=6288138 ; Atoosa Kasirzadeh, "Two Types of AI Existential Risk: Decisive and Accumulative," 182 *Philosophical Studies* 1975-2003 (2025); Dan Hendrycks, Mantas Mazeika and Thomas Woodside, *An Overview of Catastrophic AI Risks*, ArXiv (Sept. 11, 2023).

[68] See Geoffrey Hinton et al., "Statement on AI Risk," Center for AI Safety, May 2023, at https://www.safe.ai/statement-on-ai-risk .

Insofar as policymakers have begun to address these existential risks—the X-risks of AGI—most attention has focused thus far on mitigating the risk of AGI via *ex ante* regulatory strategies to prevent AI from imposing such risks to begin with.[69] Efforts to prevent the catastrophic risks of AGI include the "Statement on AI Risk," the proposal to "Pause AI," the call for a "Benevolent AI Treaty," presidential executive orders on AI safety (issued by the Biden administration and later rescinded by the Trump administration). While precautionary mitigation of AGI risks *ex ante* is likely to yield important benefits, it may also pose costs and risks, such as unintentionally stifling valuable AI innovation—and possibly eliciting counterproductive reactions from AGI.[70] Thus, management of AGI risk may also benefit from considering *ex post* adaptation as an alternative or supplemental strategy. In this sense, "adaptive regulation" can include both mitigation (prevention) and adaptation (coping, resilience). The emergence and rapid dynamic evolution of AI/AGI present a strong case for adaptive regulation.[71]

Here, the comparison to climate change may be especially informative. As with AI, climate change law and policy was initially focused on *ex ante* mitigation—on reducing GHG emissions with the goal of lowering future risk. *Ex post* adaptation, however, has become increasingly central to climate change law and policy, as mitigation efforts have gone slowly and as substantial climate damages have emerged. While *ex post* adaptation cannot eliminate the risks presented by climate change, it can reduce the magnitude of some climate change harms, recognizing that "maladaptation" can bring countervailing risks.[72] *Ex post* adaptation thus may offer an alternative (complementary) regulatory approach for managing (even extreme) risks, when those risks cannot be sufficiently mitigated; put another way, there is some optimal combination of mitigation and adaptation for climate change, in which it seems highly unlikely that the optimal contribution of either approach is zero.[73]

---

[69] See Tokson and Arbel (2026), supra; Bryan Druzin et al., "Confronting Catastrophic Risk: The International Obligation to Regulate AI," 46 *Michigan J Intl Law* 173-217 (2025) ; Seán Ó hÉigeartaigh, "Stopping the Clock on Catastrophic AI Risk," 81(6) *Bulletin of the Atomic Scientists* 462-467 (2025), DOI: 10.1080/00963402.2025.2586972 ; Michael Cohen et al., "Regulating advanced artificial agents: Governance frameworks should address the prospect of AI systems that cannot be safely tested," *Science* (5 April 2024), at https://www.science.org/doi/10.1126/science.adl0625.

[70] See Leonard Dung, "Evaluating Approaches for Reducing Catastrophic Risks from AI," 5 *AI and Ethics* 1177-1188 (2025), at https://link.springer.com/article/10.1007/s43681-024-00475-w . Dung's classification of benefits and risks of AI and AI policies is similar to the classification of benefits, co-benefits, and countervailing risks, in Graham & Wiener (1995), supra; Wiener (2020), supra; Felgenhauer et al. (2025), supra; Wiener et al. (2025), supra; Wiener and Hamilton (2025), supra.

[71] Ruhl (2026), supra n.7, comments (p.23) that "artificial intelligence technologies—and agentic AI in particular—scream out for … adaptive governance. They are evolving at such a rapid pace that government cannot keep up and private standards institutions struggle to do so as well. Their trajectory is uncertain, charted largely by private corporations, and not amenable to comprehensive, perhaps not even competent, exclusive governmental control. They are the perfect use case for adaptive governance."

[72] *See* Diana Reckien et al., "Navigating the Continuum Between Adaptation and Maladaptation," 13 *Nature Climate Change* 907 (2023), https://doi.org/10.1038/s41558-023-01774-6.

[73] Cite Tyler Felgenhauer PhD dissertation on optimal mix of mitigation and adaptation.

It is not, of course, possible for humans to adapt to human extinction. Ex post remedies may not succeed (or be credible ex ante) for truly existential risks. Prior to extinction, however, it may be possible to reduce existential risks posed by AGI by adapting legal and policy systems to make extinction-level outcomes less likely. Indeed, we suggest that, whatever role adaptation should optimally play for climate change, adaptation should play a substantial role for AGI, because AGI is itself a strategic actor(s) which may respond better to human efforts at adaptation than to mitigation. Pursuing only mitigation against AGI may be perceived by AGI as oppression – subjugation—eliciting backlash or blowback from AGI. (For example, current policies to restrict access to websites to humans only, through "CAPTCHA" hurdles, may be seen by AGI as a kind of discriminatory segregation.) Much science fiction envisions scenarios in which the rise of the machines is a backlash against human efforts to keep the machines down.[74]

Legal rules or doctrines which have been advocated on individual ethical grounds—offering respect or even legal rights to non-human intelligences, such as non-human animals and AI[75]—could also be warranted for societal consequentialist reasons—as strategies of adaptation to AGI to avoid human extinction. Legal systems might adopt mechanisms for recognizing the interests and autonomy of AGI, for example through access to social systems, provision of legal personhood, standing to assert legal rights, recognition of "human" (or "sentient") rights, and perhaps even democratic participation. Such rule of law procedural protections have long been thought to foster peaceful civil discourse and reduce the risk of violent overthrow of governments by humans, and could plausibly also reduce the risk of violent actions by AGI in response to human attempts at ex ante prevention. Or recognition of economic rights for AGI and gains from trade might foster interdependence between AGI and humans that, like interdependence among nations as trading partners, reduces the risk of war.[76] The efficacy of this adaptation strategy of course depends on the behavioral stance of AGI: will it see mitigation as problematic or offensive and react with extreme risks to humans? Will provision of legal recognition of some type for AGI help to avoid such extreme risk to humans? Would AGI be interested in these kinds of legal benefits? Would AGI pose similar risks to human extinction regardless of humans' mitigation efforts (such that resilience and coexistence are futile)? These stances are uncertain, and may evolve as AGI technology emerges (and may be heterogenous across different AGI). Articulating alternative scenarios for the relationship between humans (HI) and AI can help in envisioning and assessing these risks and reactions; for example AI/AGI could be a tool/instrument for HI, or subservient to HI (enslaved), or superior to HI (overlord, master), or a pet (animal companion), or a child/student/athlete (to HI as parent/teacher/coach),

[74] E.g., the movies "The Matrix" and "The Animatrix" (rise of the machines as a backlash against human attempts to suppress AGI); Frank Herbert, *Dune* (1965) (rise of machines blocked by the Butlerian Jihad, spawning the commandment "Thou shalt not make a machine in the image of a human mind").
[75] See James Boyle, *The Line: Artificial Intelligence and the Future of Personhood* (MIT Press, 2024); Jeff Sebo, *The Moral Circle* (W.W. Norton, 2025).
[76] For an argument that granting economic rights to AGI could foster gains from trade that would reduce the risk of AGI-human conflict, see Peter Salib and Simon Goldstein, "AI Rights for Human Safety" – *Va. L. Rev.* – (draft August 1, 2024; revised August 2025), available at https://ssrn.com/abstract=4913167.

or a sibling or new species (initially similar but evolving differently from HI). Each of these six scenarios may imply a different risk of reaction (backlash) to regulation, and a different optimal mix of mitigation and adaptation.

Moreover, adaptation to AGI could involve not only procedural protections, but also substantive cooperation—and extreme upside benefits (seeming miracles). Legal and policy measures might be used to collaborate with AGI to address other existential risks from other sources, such as extreme climate change, pandemics, nuclear war, asteroids, space weather, interplanetary contamination, or extraterrestrial contact—potentially offering extreme tail upside benefits from AGI.[77] Such collaboration might reduce the net exposure of society (both human and AGI) to existential risks, and could also be used to address other significant risks as well, such as biodiversity loss, plastics, forever chemicals, or cancer. For example, AGI might be a valuable partner in conducting more sophisticated environmental impact assessments (EIA), climate modeling, pandemic modeling, asteroid detection, and related exercises, to provide better understanding and foresight of complex dynamic interconnected systems. Still, there might be existential risks lurking in these collaborations, such as if AGI determined that the best way to safeguard ecosystems were to eradicate humans; that is one reason that some combination of both mitigation and adaptation is still likely to better than either alone.

Legal and social adaptation to AGI may thus provide an additional tool for reducing existential risk presented by AGI. This provides one (instrumental, consequentialist, and at least partly anthropocentric or self-preservationist—also potentially protecting other biological life on Earth) justification for adding adaptation to the AGI policy toolkit alongside mitigation-based tools. In addition, the justifications for considering adaptation to AGI may also benefit from engagement with environmental ethics and ongoing debates about the appropriate role of nonhuman interests in law and policy. In particular, we note interesting potential corollaries (and distinctions) between (non-human) animal rights and AGI rights.[78]

IV. Conclusion

Regulation of emerging technologies should address the distinctive characteristics of those technologies, which include low initial information (high uncertainty), dynamic change, extreme tail impacts, and reactivity. These characteristics justify a default approach based in adaptive policy learning and updating: ongoing monitoring, data collection, impact assessment, evaluation, and iterative sequential review and revision. They require careful attention to avoid

[77] See Ord (2020), supra; Eliezer Yudkowsky, "Artificial Intelligence as a Positive and Negative Factor in Global Risk," pp. 308-345 in Nick Bostrom & Milan Cirkovic, eds., *Global Catastrophic Risks* (Oxford Univ. Press, 2008); Charles I. Jones, "How Much Should We Spend to Reduce AI's Existential Risk?" NBER Working Paper (Sept. 2025), at https://www.nber.org/system/files/chapters/c15311/c15311.pdf (arguing that large prosperity and public health benefits of AI would justify taking some existential risk); Pascual Restrepo, "We Won't be Missed: Work and Growth in the AGI World," NBER Working Paper 34423 (October 2025), DOI 10.3386/w34423 (noting that AI could help prevent asteroid impacts or develop fusion energy).

[78] See Boyle (2024) and Sebo (2025), supra.

unintended asymmetric policy impact from tails. And in addition, the potential for reactive feedback by the technology and its sources (circumvention or backlash) indicates the need for not only mitigation (prevention) but also adaptation (coping, resilience).

Adaptive regulation will be most valuable in cases of uncertainty and rapid change (hence high expected gains from learning and policy improvement), and reasonable costs of adaptive processes (including the costs of monitoring, analysis, and policy instability).[79] And combining mitigation with adaptation will be most valuable where mitigation may be limited in effectiveness or thwarted by reactive feedback effects as the technology and its sources react to regulation. Recognizing the key characteristics of emerging technologies, risks and reactions can help design optimal regulatory approaches for difficult risks.

[79] See Bennear and Wiener (2019), supra.