International Center for Law & Economics

Comments of the International Center for Law & Economics, Dual Use Foundation Artificial Intelligence Models with Widely Available Model Weights

Docket No. NTIA–240216-0052

March 27, 2024

Authored by:

Kristian Stout (Director of Innovation Policy, International Center for Law & Economics)

I. Introduction

We thank the National Telecommunications and Information Administration (NTIA) for the opportunity to contribute to this request for comments (RFC) in the "Dual Use Foundation Artificial Intelligence Models with Widely Available Model Weights" proceeding. In these comments, we endeavor to offer recommendations to foster the innovative and responsible production of artificial intelligence (AI), encompassing both open-source and proprietary models. Our comments are guided by a belief in the transformative potential of AI, while recognizing NTIA's critical role in guiding the development of regulations that not only protect consumers but also enable this dynamic field to flourish. The agency should seek to champion a balanced and forward-looking approach toward AI technologies that allows them to evolve in ways that maximize their social benefits, while navigating the complexities and challenges inherent in their deployment.

NTIA's question "How should [the] potentially competing interests of innovation, competition, and security be addressed or balanced?"[1] gets to the heart of ongoing debates about AI regulation. There is no panacea to be discovered, as all regulatory choices require balancing tradeoffs. It is crucial to bear this in mind when evaluating, *e.g.*, regulatory proposals that implicitly treat AI as inherently dangerous and regard as obvious that stringent regulation is the only effective strategy to mitigate such risks.[2] Such presumptions discount AI's unknown but potentially enormous capacity to produce innovation, and inadequately account for other tradeoffs inherent to imposing a risk-based framework (*e.g.*, requiring disclosure of trade secrets or particular kinds of transparency that could yield new cybersecurity attack vectors). Adopting an overly cautious stance risks not only stifling AI's evolution, but may also preclude a fulsome exploration of its potential to foster social, economic, and technological advancement. A more restrictive regulatory environment may also render AI technologies more homogenous and smother development of the kinds of diverse AI applications needed to foster robust competition and innovation.

We observe this problematic framing in the executive order (EO) that serves as the provenance of this RFC.[3] The EO repeatedly proclaims the importance of "[t]he responsible development

---

[1] *Dual Use Foundation Artificial Intelligence Models With Widely Available Model Weights*, Docket No. 240216-0052, 89 FR 14059, NATIONAL TELECOMMUNICATIONS AND INFORMATION ADMINISTRATION (Mar. 27, 2024) at 14063, question 8(a) [hereinafter "RFC"].

[2] *See, e.g.*, Kristian Stout, *Systemic Risk and Copyright in the EU AI Act*, TRUTH ON THE MARKET (Mar. 19, 2024), https://truthonthemarket.com/2024/03/19/systemic-risk-and-copyright-in-the-eu-ai-act.

[3] Exec. Order No. 14110, 88 F.R. 75191 (2023), https://www.federalregister.gov/documents/2023/11/01/2023-24283/safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence?_fsi=C0CdBzzA [hereinafter "EO"].

and use of AI" in order to "mitigate[e] its substantial risks."[4] Specifically, the order highlights concerns over "dual-use foundation models"—*i.e.*, AI systems that, while beneficial, could pose serious risks to national security, national economic security, national public health, or public safety.[5] Concerningly, one of the categories the EO flags as illicit "dual use" are systems "permitting the evasion of human control or oversight through means of deception or obfuscation."[6] This open-ended category could be interpreted so broadly that essentially any general-purpose generative-AI system would classify.

The EO also repeatedly distinguishes "open" versus "closed" approaches to AI development, while calling for "responsible" innovation and competition.[7] On our reading, the emphasis the EO places on this distinction raises alarm bells about the administration's inclination to stifle innovation through overly prescriptive regulatory frameworks, diminishment of the intellectual property rights that offer incentives for innovation, and regulatory capture that favors incumbents over new entrants. In favoring one model of AI development over another, the EO's prescriptions could inadvertently hamper the dynamic competitive processes that are crucial both for technological progress and for the discovery of solutions to the challenges that AI technology poses.

Given the inchoate nature of AI technology—much less the uncertain markets in which that technology will ultimately be deployed and commercialized—NTIA has an important role to play in elucidating for policymakers the nuances that might lead innovators to choose an open or closed development model, without presuming that one model is inherently better than the other—or that either is necessarily "dangerous." Ultimately, the preponderance of AI risks will almost certainly emerge idiosyncratically. It will be incumbent on policymakers to address such risks in an iterative fashion as they become apparent. For now, it is critical to resist the urge to enshrine crude and blunt categories for the heterogeneous suite of technologies currently gathered under the broad banner of "AI."

Section II of these comments highlights the importance of grounding AI regulation in actual harms, rather than speculative risks, while outlining the diversity of existing AI technologies and the need for tailored approaches. Section III starts with discussion of some of the benefits and challenges posed by both open and closed approaches to AI development, while cautioning against overly prescriptive definitions of "openness" and advocating flexibility in regulatory frameworks. It proceeds to examine the EO's prescription to regulate so-called "dual-use" foundation models, underscoring some potential unintended consequences for open-source

[4] *See*, *e.g.*, EO at §§ 1; 2(c), 5.2(e)(ii); and § 8(c);

[5] *Id.* at § 3(k).

[6] *Id.* at § (k)(iii).

[7] *Id.* at § 4.6. As NTIA notes, the administration refers to "widely available model weight," which is equivalent to "open foundation models" in this proceeding. RFC at 14060.

AI development and international collaboration. Section IV offers some principles to craft an effective regulatory model for AI, including distinguishing between low-risk and high-risk applications, avoiding static regulatory approaches, and adopting adaptive mechanisms like regulatory sandboxes and iterative rulemaking. Section V concludes.

II. Risk Versus Harm in AI Regulation

In many of the debates surrounding AI regulation, disproportionate focus is placed on the need to mitigate risks, without sufficient consideration of the immense benefits that AI technologies could yield. Moreover, because these putative risks remain largely hypothetical, proposals to regulate AI descend quickly into an exercise in shadowboxing.

Indeed, there is no single coherent definition of what even constitutes "AI." The term encompasses a wide array of technologies, methodologies, and applications, each with distinct characteristics, capabilities, and implications for society. From foundational models that can generate human-like text, to algorithms capable of diagnosing diseases with greater accuracy than human doctors, to "simple" algorithms that facilitate a more tailored online experience, AI applications and their underlying technologies are as varied as they are transformative.

This diversity has profound implications for the regulation and development of AI. Very different regulatory considerations are relevant to AI systems designed for autonomous vehicles than for those used in financial algorithms or creative-content generation. Each application domain comes with its own set of risks, benefits, ethical dilemmas, and potential social impacts, necessitating tailored approaches to each use case. And none of these properties of AI map clearly onto the "open" and "closed" designations highlighted by the EO and this RFC. This counsels for focus on specific domains and specific harms, rather than how such technologies are developed.[8]

As in prior episodes of fast-evolving technologies, what is considered cutting-edge AI today may be obsolete tomorrow. This rapid pace of innovation further complicates the task of crafting policies and regulations that will be both effective and enduring. Policymakers and regulators must navigate this terrain with a nuanced understanding of AI's multifaceted nature, including by embracing flexible and adaptive regulatory frameworks that can accommodate AI's continuing evolution.[9] A one-size-fits-all approach could inadvertently stifle innovation or entrench the dominance of a few large players by imposing barriers that disproportionately affect smaller entities or emerging technologies.

[8] For more on the "open" vs "closed" distinction and its poor fit as a regulatory lens, *see*, *infra*, at nn. 18-39 and accompanying text.

[9] Adaptive regulatory frameworks are discussed, *infra*, at nn. 40-51 and accompanying text.

Experts in law and economics have long scrutinized both market conduct and regulatory rent seeking that serve to enhance or consolidate market power by disadvantaging competitors, particularly through increasing the costs incurred by rivals.[10] Various tactics may be employed to undermine competitors or exclude them from the market that do not involve direct price competition. It is widely recognized that "engaging with legislative bodies or regulatory authorities to enact regulations that negatively impact competitors" produces analogous outcomes.[11] It is therefore critical that the emerging markets for AI technologies not engender opportunities for firms to acquire regulatory leverage over rivals. Instead, recognizing the plurality of AI technologies and encouraging a multitude of approaches to AI development could help to cultivate a more vibrant and competitive ecosystem, driving technological progress forward and maximizing AI's potential social benefits.

This overarching approach counsels skepticism about risk-based regulatory frameworks that fail to acknowledge how the theoretical harms of one type of AI system may be entirely different from those of another. Obviously, the regulation of autonomous drones is a very different sort of problem than the regulation of predictive policing or automated homework tutors. Even within a single circumscribed domain of generative AI—such as "smart chatbots" like ChatGPT or Claude—different applications may present entirely different kinds of challenges. A highly purpose-built version of such a system might be employed by government researchers to develop new materiel for the U.S. Armed Forces, while a general-purpose commercial chatbot would employ layers of protection to ensure that ordinary users couldn't learn how to make advanced weaponry. Rather treating "chatbots" as possible vectors for weapons development, a more appropriate focus would target high-capability systems designed to assist in developing such systems. Were it the case that a general-purpose chatbot inadvertently revealed some information on building weapons, all incentives would direct that AI's creators to treat that as a bug to fix, not a feature to expand.

Take, for example, the recent public response to the much less problematic AI-system malfunctions that accompanied Google's release of its Gemini program.[12] Gemini was found to generate historically inaccurate images, such as ethnically diverse U.S. senators from the 1800s, including women.[13] Google quickly acknowledged that it did not intend for Gemini to create inaccurate historical images and turned off the image-generation feature to allow time

---

[10] *See* Steven C. Salop & David T. Scheffman, *Raising Rivals' Costs*, 73:2 AM. ECON. R. 267, 267–71 (1983), http://www.jstor.org/stable/1816853.

[11] *See* Steven C. Salop & David T. Scheffman, *Cost-Raising Strategies*, 36:1 J. INDUS. ECON. 19 (1987), https://doi.org/10.2307/2098594.

[12] Cindy Gordon, *Google Pauses Gemini AI Model After Latest Debacle*, FORBES (Feb. 29, 2024), https://www.forbes.com/sites/cindygordon/2024/02/29/google-latest-debacle-has-paused-gemini-ai-model/?sh=3114d093536c.

[13] *Id.*

for the company to work on significant improvements before re-enabling it.[14] While Google blundered in its initial release, it had every incentive to discover and remedy the problem. The market response provided further incentive for Google to get it right in the future.[15] Placing the development of such systems under regulatory scrutiny because some users *might* be able to jailbreak a model and generate *some* undesirable material would create disincentives to the production of AI systems more generally, with little gained in terms of public safety.

Rather than focus on the speculative risks of AI, it is essential to ground regulation in the need to address tangible harms that stem from the observed impacts of AI technologies on society. Moreover, focusing on realistic harms would facilitate a more dynamic and responsive regulatory approach. As AI technologies evolve and new applications emerge, so too will the potential harms. A regulatory framework that prioritizes actual harms can adapt more readily to these changes, enabling regulators to update or modify policies in response to new evidence or social impacts. This flexibility is particularly important for a field like AI, where technological advancements could quickly outpace regulation, creating gaps in oversight that may leave individuals and communities vulnerable to harm.

Furthermore, like any other body of regulatory law, AI regulation must be grounded in empirical evidence and data-driven decision making. Demanding a solid evidentiary basis as a threshold for intervention would help policymakers to avoid the pitfalls of reacting to sensationalized or unfounded AI fears. This would not only enhance regulators' credibility with stakeholders, but would also ensure that resources are dedicated to addressing the most pressing and substantial issues arising from the development of AI.

III. The Regulation of Foundation Models

NTIA is right to highlight the tremendous promise that attends the open development of AI technologies:

> Dual use foundation models with widely available weights (referred to here as open foundation models) could play a key role in fostering growth among less resourced actors, helping to widely share access to AI's benefits.... Open foundation models can be readily adapted and fine-tuned to specific tasks and possibly make it easier for system developers to scrutinize the role foundation models play in larger AI systems, which is important for rights- and safety-impacting AI systems (e.g. healthcare, education, housing, criminal justice, online platforms etc.)
>
> ...Historically, widely available programming libraries have given researchers the ability to simultaneously run and understand algorithms created by other

[14] *Id.*

[15] Breck Dumas, *Google Loses $96B in Value on Gemini Fallout as CEO Does Damage Control*, YAHOO FINANCE (Feb. 28, 2024), https://finance.yahoo.com/news/google-loses-96b-value-gemini-233110640.html.

> programmers. Researchers and journals have supported the movement towards open science, which includes sharing research artifacts like the data and code required to reproduce results.[16]

The RFC proceeds to seek input on how to define "open" and "widely available."[17] These, however, are the wrong questions. NTIA should instead proceed from the assumption that there are no harms inherent to either "open" or "closed" development models; it should be seeking input on anything that might give rise to discrete harms in *either* open or closed systems.

NTIA can play a valuable role by recommending useful alterations to existing law where gaps currently exist, regardless of the business or distribution model employed by the AI developer. In short, there is nothing necessarily more or less harmful about adopting an "open" or a "closed" approach to software systems. The decision to pursue one path over the other will be made based on the relevant tradeoffs that particular firms face. Embedding such distinctions in regulation is arbitrary, at best, and counterproductive to the fruitful development of AI, at worst.

A. 'Open' or 'Widely Available' Model Weights

To the extent that NTIA is committed to drawing distinctions between "open" and "closed" approaches to developing foundation models, it should avoid overly prescriptive definitions of what constitutes "open" or "widely available" model weights that could significantly hamper the progress and utility of AI technologies.

Imposing narrow definitions risks creating artificial boundaries that fail to accurately reflect AI's technical and operational realities. They could also inadvertently exclude or marginalize innovative AI models that fall outside those rigid parameters, despite their potential to contribute positively to technological advancement and social well-being. For instance, a definition of "open" that requires complete public accessibility without any form of control or restriction might discourage organizations from sharing their models, fearing misuse or loss of intellectual property.

Moreover, prescriptive definitions could stifle the organic growth and evolution of AI technologies. The AI field is characterized by its rapid pace of change, where today's cutting-edge models may become tomorrow's basic tools. Prescribing fixed criteria for what constitutes "openness" or "widely available" risks anchoring the regulatory landscape to this specific moment in time, leaving the regulatory framework less able to adapt to future developments and innovations.

---

[16] RFC at 14060.

[17] RFC at 14062, question 1.

Given AI developers' vast array of applications, methodologies, and goals, it is imperative that any definitions of "open" or "widely available" model weights embrace flexibility. A flexible approach would acknowledge how the various stakeholders within the AI ecosystem have differing needs, resources, and objectives, from individual developers and academic researchers to startups and large enterprises. A one-size-fits-all definition of "openness" would fail to accommodate this diversity, potentially privileging certain forms of innovation over others and skewing the development of AI technologies in ways that may not align with broader social needs.

Moreover, flexibility in defining "open" and "widely available" must allow for nuanced understandings of accessibility and control. There can, for example, be legitimate reasons to limit openness, such as protecting sensitive data, ensuring security, and respecting intellectual-property rights, while still promoting a culture of collaboration and knowledge sharing. A flexible regulatory approach would seek a balanced ecosystem where the benefits of open AI models are maximized, and potential risks are managed effectively.

B. The Benefits of 'Open' vs 'Closed' Business Models

NTIA asks:

> What benefits do open model weights offer for competition and innovation, both in the AI marketplace and in other areas of the economy? In what ways can open dual-use foundation models enable or enhance scientific research, as well as education/training in computer science and related fields?[18]

An open approach to AI development has obvious benefits, as NTIA has itself acknowledged in other contexts.[19] Open-foundation AI models represent a transformative force, characterized by their accessibility, adaptability, and potential for widespread application across various sectors. The openness of these models may serve to foster an environment conducive to innovation, wherein developers, researchers, and entrepreneurs can build on existing technologies to create novel solutions tailored to diverse needs and challenges.

The inherent flexibility of open-foundation models can also catalyze a competitive market, encouraging a healthy ecosystem where entities ranging from startups to established corporations may all participate on roughly equal footing. By lowering some entry barriers related to access to basic AI technologies, this competitive environment can further drive

[18] RFC at 14062, question 3(a).

[19] Department of Commerce, *Competition in the Mobile Application Ecosystem* (2023), https://www.ntia.gov/report/2023/competition-mobile-app-ecosystem ("While retaining appropriate latitude for legitimate privacy, security, and safety measures, Congress should enact laws and relevant agencies should consider measures (such as rulemaking) designed to open up distribution of lawful apps, by prohibiting… barriers to the direct downloading of applications.").

technological advancements and price efficiencies, ultimately benefiting consumers and society at-large.

But more "closed" approaches can also prove very valuable. As NTIA notes in this RFC, it is rarely the case that a firm pursues a purely open or closed approach. These terms exist along a continuum, and firms blend models as necessary.[20] And just as firms readily mix elements of open and closed business models, a regulator should be agnostic about the precise mix that firms employ, which ultimately must align with the realities of market dynamics and consumer preferences.

Both open and closed approaches offer distinct benefits and potential challenges. For instance, open approaches might excel in fostering a broad and diverse ecosystem of applications, thereby appealing to users and developers who value customization and variety. They can also facilitate a more rapid dissemination of innovation, as they typically impose fewer restrictions on the development and distribution of new applications. Conversely, closed approaches, with their curated ecosystems, often provide enhanced security, privacy, and a more streamlined user experience. This can be particularly attractive to users less inclined to navigate the complexities of open systems. Under the right conditions, closed systems can likewise foster a healthy ecosystem of complementary products.

The experience of modern digital platforms demonstrates that there is no universally optimal approach to structuring business activities, thus illustrating the tradeoffs inherent in choosing among open and closed business models. The optimal choice depends on the specific needs and preferences of the relevant market participants. As Jonathan M. Barnett has noted:

> Open systems may yield no net social gain over closed systems, can pose a net social loss under certain circumstances, and . . . can impose a net social gain under yet other circumstances.[21]

Similar considerations apply in the realm of AI development. Closed or semi-closed ecosystems can offer such advantages as enhanced security and curated offerings, which may appeal to certain users and developers. These benefits, however, may come at the cost of potentially limited innovation, as a firm must rely on its own internal processes for research and development. Open models, on the other hand, while fostering greater collaboration and creativity, may also introduce risks related to quality control, intellectual-property protection, and a host of other concerns that may be better controlled in a closed business model. Even along innovation dimensions, closed platforms can in many cases outperform open models.

---

[20] RFC at 14061 ("'openness' or 'wide availability' of model weights are also terms without clear definition or consensus. There are gradients of 'openness,' ranging from fully 'closed' to fully 'open'").

[21] *See* Jonathan M. Barnett, *The Host's Dilemma: Strategic Forfeiture in Platform Markets for Informational Goods*, 124 HARV. L. REV. 1861, 1927 (2011).

With respect to digital platforms like the App Store and Google Play Store, there is a "fundamental welfare tradeoff between two-sided proprietary…platforms and two-sided platforms which allow 'free entry' on both sides of the market."[22] Consequently, "it is by no means obvious which type of platform will create higher product variety, consumer adoption and total social welfare."[23]

To take another example, consider the persistently low adoption rates for consumer versions of the open-source Linux operating system, versus more popular alternatives like Windows or MacOS.[24] A closed model like Apple's MacOS is able to outcompete open solutions by better leveraging network effects and developing a close relationship with end users.[25] Even in this example, adoption of open versus closed models varies across user types, with, *e.g.*, developers showing a strong preference for Linux over Mac, and only a slight preference for Windows over Linux.[26] This underscores the point that the suitability of an open or closed model varies not only by firm and product, nor even solely by user, but by the unique fit of a particular model for a particular user in a particular context. Many of those Linux-using developers will likely *not* use it on their home computing device, for example, even if they prefer it for work.

The dynamics among consumers and developers further complicate prevailing preferences for open or closed models. For some users, the security and quality assurance provided by closed ecosystems outweigh the benefits of open systems' flexibility. On the developer side, the lower barriers to entry in more controlled ecosystems that smooth the transaction costs associated with developing and marketing applications can democratize application development, potentially leading to greater innovation within those ecosystems. Moreover, distinctions between open and closed models can play a critical role in shaping inter-brand competition. A regulator placing its thumb on the business-model scale would push the relevant markets toward less choice and lower overall welfare.[27]

By differentiating themselves through a focus on ease-of-use, quality, security, and user experience, closed systems contribute to a vibrant competitive landscape where consumers have clear choices between differing "brands" of AI. Forcing an AI developer to adopt practices that

---

[22] *Id.* at 2.

[23] *Id.* at 3.

[24] *Desktop Operating System Market Share Worldwide Feb 2023 - Feb 2024*, STATCOUNTER, https://gs.statcounter.com/os-market-share/desktop/worldwide (*last visited* Mar. 27, 2024).

[25] Andrei Hagiu, *Proprietary vs. Open Two-Sided Platforms and Social Efficiency* (HARV. BUS. SCH. STRATEGY UNIT, Working Paper No. 09-113, 2006).

[26] Joey Sneddon, *More Developers Use Linux than Mac, Report Shows*, OMG LINUX (Dec. 28, 2022), https://www.omglinux.com/devs-prefer-linux-to-mac-stackoverflow-survey.

[27] *See* Michael L. Katz & Carl Shapiro, *Systems Competition and Network Effects*, 8 J. ECON. PERSP. 93, 110 (1994), ("[T]he primary cost of standardization is loss of variety: consumers have fewer differentiated products to pick from, especially if standardization prevents the development of promising but unique and incompatible new systems").

align with a regulator's preconceptions about the relative value of "open" and "closed" risks homogenizing the market and diminishing the very competition that spurs innovation and consumer choice.

Consider some of the practical benefits sought by deployers when choosing between open and closed models. For example, it's not straightforward to say close is inherently better than open when considering issues of data sharing or security; even here, there are tradeoffs. Open innovation in AI—characterized by the sharing of data, algorithms, and methodologies within the research community and beyond—can mitigate many of the risks associated with model development. This openness fosters a culture of transparency and accountability, where AI models and their applications are subject to scrutiny by a broad community of experts, practitioners, and the general public. This collective oversight can help to identify and address potential safety and security concerns early in the development process, thus enhancing AI technologies' overall trustworthiness.

By contrast, a closed system may implement and enforce standardized security protocols more quickly. A closed system may have a sharper, more centralized focus on providing data security to users, which may perform better along some dimensions. And while the availability of code may provide security in some contexts, in other circumstances, closed systems perform better.[28]

In considering ethical AI development, different types of firms should be free to experiment with different approaches, even blending them where appropriate. For example, Claude's approach to "Collective Constitutional AI" adopts what is arguably a "semi-open" model, blending proprietary elements with certain aspects of openness to foster innovation, while also maintaining a level of control.[29] This model might strike an appropriate balance, in that it ensures some degree of proprietary innovation and competitive advantage while still benefiting from community feedback and collaboration.

On the other hand, fully open-source development could lead to a different, potentially superior result that meets a broader set of needs through community-driven evolution and iteration. There is no way to determine, *ex ante*, that either an open or a closed approach to AI development will inherently provide superior results for developing "ethical" AI. Each has its place, and, most likely, the optimal solutions will involve elements of both approaches.

In essence, codifying a regulatory preference for one business model over the other would oversimplify the intricate balance of tradeoffs inherent to platform ecosystems. Economic

---

[28] *See. e.g.*, Nokia, *Threat Intelligence Report 2020* (2020), https://www.nokia.com/networks/portfolio/cyber-security/threat-intelligence-report-2020; Randal C. Picker, *Security Competition and App Stores*, NETWORK LAW REVIEW (Aug. 23, 2021), https://www.networklawreview.org/picker-app-stores.

[29] *Collective Constitutional AI: Aligning a Language Model with Public Input*, ANTHROPIC (Oct. 17, 2023), https://www.anthropic.com/news/collective-constitutional-ai-aligning-a-language-model-with-public-input.

theory and empirical evidence suggest that both open and closed platforms can drive innovation, serve consumer interests, and stimulate healthy competition, with all of these considerations depending heavily on context. Regulators should therefore aim for flexible policies that support coexistence of diverse business models, fostering an environment where innovation can thrive across the continuum of openness.

C. Dual-Use Foundation Models and Transparency Requirements

The EO and the RFC both focus extensively on so-called "dual-use" foundation models:

> Foundation models are typically defined as, "powerful models that can be fine-tuned and used for multiple purposes." Under the Executive Order, a "dual-use foundation model" is "an AI model that is trained on broad data; generally uses self-supervision, contains at least tens of billions of parameters; is applicable across a wide range of contexts; and that exhibits, or could be easily modified to exhibit, high levels of performance at tasks that pose a serious risk to security, national economic security, national public health or safety, or any combination of those matters...."[30]

But this framing will likely do more harm than good. As noted above, the terms "AI" or "AI model" are frequently invoked to refer to very different types of systems. Further defining these models as "dual use" is also unhelpful, as virtually *any* tool in existence can be "dual use" in this sense. Certainly, from a certain perspective, all software—particularly highly automated software—can pose a serious risk to "national security" or "safety." Encryption and other privacy-protecting tools certainly fit this definition.[31] While it is crucial to mitigate harms associated with the misuse of AI technologies, the blanket treatment of all foundation models under this category is overly simplistic.

The EO identifies certain clear risks, such as the possibility that models could aid in the creation of chemical, biological, or nuclear weaponry. These categories are obvious subjects for regulatory control, but the EO then appears to open a giant definitional loophole that threatens to subsume virtually any useful AI system. It employs expansive terminology to describe a more generalized threat—specifically, that dual-use models could "[permit] the evasion of human control or oversight through means of deception or obfuscation."[32] Such language could encompass a wide array of general-purpose AI models. Furthermore, by labeling

---

[30] RFC at 14061.

[31] *Encryption and the "Going Dark" Debate*, CONGRESSIONAL RESEARCH SERVICE (2017), https://crsreports.congress.gov/product/pdf/R/R44481.

[32] EO at. § 3(k)(iii).

systems capable of bypassing human decision making as "dual use," the order implicitly suggests that all AI could pose such risk as warrants national-security levels of scrutiny.

Given the EO's broad definition of AI as "a machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations, or decisions influencing real or virtual environments," numerous software systems not typically even considered AI might be categorized as "dual-use" models.[33] Essentially, any sufficiently sophisticated statistical-analysis tool could qualify under this definition.

A significant repercussion of the EO's very broad reporting mandates for dual-use systems, and one directly relevant to the RFC's interest in promoting openness, is that these might chill open-source AI development.[34] Firms dabbling in AI technologies—many of which might not consider their projects to be dual use—might keep their initiatives secret until they are significantly advanced. Faced with the financial burden of adhering to the EO's reporting obligations, companies that lack a sufficiently robust revenue model to cover both development costs and legal compliance might be motivated to dodge regulatory scrutiny in the initial phases, consequently dampening the prospects for transparency.

It is hard to imagine how open-source AI projects could survive in such an environment. Open-source AI code libraries like TensorFlow[35] and PyTorch[36] foster remarkable innovation by allowing developers to create new applications that use cutting-edge models. How could a paradigmatic startup developer working out of a garage genuinely commit to open-source development if tools like these fall under the EO's jurisdiction? Restricting access to the weights that models use—let alone avoiding open-source development entirely—may hinder independent researchers' ability to advance the forefront of AI technology.

Moreover, scientific endeavors typically benefit from the contributions of researchers worldwide, as collaborative efforts on a global scale are known to fast-track innovation. The pressure the EO applies to open-source development of AI tools could curtail international cooperation, thereby distancing American researchers from crucial insights and collaborations. For example, AI's capacity to propel progress in numerous scientific areas is potentially vast—*e.g.*, utilizing MRI images and deep learning for brain-tumor diagnoses[37] or employing machine

---

[33] EO at § 3(b).

[34] EO at § 4.2 (requiring companies developing dual-use foundation models to provide ongoing reports to the federal government on their activities, security measures, model weights, and red-team testing results).

[35] *An End-to-End Platform for Machine Learning*, TENSORFLOW, https://www.tensorflow.org (*last visited* Mar. 27, 2024).

[36] *Learn the Basics*, PYTORCH, https://pytorch.org/tutorials/beginner/basics/intro.html (*last visited* Mar. 27, 2024).

[37] Akmalbek Bobomirzaevich Abdusalomov, Mukhriddin Mukhiddinov, & Taeg Keun Whangbo, *Brain Tumor Detection Based on Deep Learning Approaches and Magnetic Resonance Imaging*, 15(16) CANCERS (BASEL) 4172 (2023), *available at* https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10453020.

learning to push the boundaries of materials science.[38] Such research does not benefit from stringent secrecy, but thrives on collaborative development. Enabling a broader community to contribute to and expand upon AI advancements supports this process.

Individuals respond to incentives. Just as how well-intentioned seatbelt laws paradoxically led to an uptick in risky driving behaviors,[39] ill-considered obligations placed on open-source AI developers could unintentionally stifle the exchange of innovative concepts crucial to maintain the United States' leadership in AI innovation.

IV. Regulatory Models that Support Innovation While Managing Risks Effectively

In the rapidly evolving landscape of artificial intelligence (AI), it is paramount to establish governance and regulatory frameworks that both encourage innovation and ensure safety and ethical integrity. An effective regulatory model for AI should be adaptive, principles-based, and foster a collaborative environment among regulators, developers, researchers, and the broader community. A number of principles can help in developing this regime.

A. Low-Risk vs High-Risk AI

First, a clear distinction should be made between low-risk AI applications that enhance operational efficiency or consumer experience and high-risk applications that could have significant safety implications. Low-risk applications like search algorithms and chatbots should be governed by a set of baseline ethical guidelines and best practices that encourage innovation, while ensuring basic standards are met. On the other hand, high-risk applications—such as those used by law enforcement or the military—would require more stringent review processes, including impact assessments, ethical reviews, and ongoing monitoring to mitigate potentially adverse effects.

Contrast this with the recently enacted AI Act in the European Union, and its decision to create presumptions of risk for general purpose AI (GPAI) systems, such as large language models (LLMs), that present what the EU has termed so-called "systemic risk."[40] Article 3(65) of the AI Act defines systemic risk as "a risk that is specific to the high-impact capabilities of general-purpose AI models, having a significant impact on the Union market due to their

---

[38] Keith T. Butler, *et al.*, *Machine Learning for Molecular and Materials Science,* 559 NATURE 547 (2018), *available at* https://www.nature.com/articles/s41586-018-0337-2.

[39] *The Peltzman Effect*, THE DECISION LAB, https://thedecisionlab.com/reference-guide/psychology/the-peltzman-effect (*last visited* Mar. 27, 2024).

[40] European Parliament, European Parliament legislative Resolution of 13 March 2024 on the Proposal for a Regulation of the European Parliament and of the Council on Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts, COM/2021/206, *available at* https://www.europarl.europa.eu/doceo/document/TA-9-2024-0138_EN.html [hereinafter "EU AI Act"].

reach, or due to actual or reasonably foreseeable negative effects on public health, safety, public security, fundamental rights, or the society as a whole, that can be propagated at scale across the value chain."[41]

This definition bears similarities to the "Hand formula" in U.S. tort law, which balances the burden of precautions against the probability and severity of potential harm to determine negligence.[42] The AI Act's notion of systemic risk, however, is applied more broadly to entire categories of AI systems based on their theoretical potential for widespread harm, rather than on a case-by-case basis.

The designation of LLMs as posing "systemic risk" is problematic for several reasons. It creates a presumption of risk merely based on a GPAI system's scale of operations, without any consideration of the actual likelihood or severity of harm in specific use cases. This could lead to unwarranted regulatory intervention and unintended consequences that hinder the development and deployment of beneficial AI technologies. And this broad definition of systemic risk gives regulators significant leeway to intervene in how firms develop and release their AI products, potentially blocking access to cutting-edge tools for European citizens, even in the absence of tangible harms.

While it is important to address potential risks associated with AI systems, the AI Act's approach risks stifling innovation and hindering the development of beneficial AI technologies within the EU.

B. Avoid Static Regulatory Approaches

AI regulators are charged with overseeing a dynamic and rapidly developing market, and should therefore avoid erecting a rigid framework that force new innovations into ill-fitting categories. The "regulatory sandbox" may provide a better model to balance innovation with risk management. By allowing developers to test and refine AI technologies in a controlled environment under regulatory oversight, sandboxes can be used to help identify and address potential issues before wider deployment, all while facilitating dialogue between innovators and regulators. This approach not only accelerates the development of safe and ethical AI solutions, but also builds mutual understanding and trust. Where possible, NTIA should facilitate policy experimentation with regulatory sandboxes in the AI context.

---

[41] *Id.* at Art. 3(65).

[42] *See* Stephen G. Gilles, *On Determining Negligence: Hand Formula Balancing, the Reasonable Person Standard, and the Jury*, 54 VANDERBILT L. REV. 813, 842-49 (2001).

Meta's Open Loop program is an example of this kind of experimentation.[43] This program is a policy prototyping research project focused on evaluating the National Institute of Standards and Technology (NIST) AI Risk Management Framework (RMF) 1.0.[44] The goal is to assess whether the framework is understandable, applicable, and effective in assisting companies to identify and manage risks associated with generative AI. It also provides companies an opportunity to familiarize themselves with the NIST AI RMF and its application in risk-management processes for generative AI systems. Additionally, it aims to collect data on existing practices and offer feedback to NIST, potentially influencing future RMF updates.

1. Regulation as a Discovery Process

Another key principle is to ensure that regulatory mechanisms are adaptive. Some examples of adaptive mechanisms are iterative rulemaking and feedback loops that allow regulations to be updated continuously in response to new developments and insights. Such mechanisms enable policymakers to respond swiftly to technological breakthroughs, ensuring that regulations remain relevant and effective, without stifling innovation.

Geoffrey Manne & Gus Hurwitz have recently proposed a framework for "regulation as a discovery process" that could be adapted to AI.[45] They argue for a view of regulation not merely as a mechanism for enforcing rules, but as a process for discovering information that can inform and improve regulatory approaches over time. This perspective is particularly pertinent to AI, where the pace of innovation and the complexity of technologies often outstrip regulators' understanding and ability to predict future developments. This framework:

> in its simplest formulation, asks regulators to consider that they might be wrong. That they might be asking the wrong questions, collecting the wrong information, analyzing it the wrong way—or even that Congress has given them the wrong authority or misunderstood the problem that Congress has tasked them to address.[46]

That is to say, an adaptive approach to regulation requires epistemic humility, with the understanding that, particularly for complex, dynamic industries:

> there is no amount of information collection or analysis that is guaranteed to be "enough." As Coase said, the problem of social cost isn't calculating what those

---

[43] *See Open Loop's First Policy Prototyping Program in the United States*, META, https://www.usprogram.openloop.org (*last visited* Mar. 27. 2024).

[44] *Id.*

[45] Justin (Gus) Hurwitz & Geoffrey A. Manne, *Pigou's Plumber: Regulation as a Discovery Process*, SSRN (2024), *available at* https://laweconcenter.org/resources/pigous-plumber.

[46] *Id.* at 32.

> costs are so that we can eliminate them, but ascertaining how much of those social costs society is willing to bear.[47]

In this sense, modern regulators' core challenge is to develop processes that allow for iterative development of knowledge, which is always in short supply. This requires a shift in how an agency conceptualizes its mission, from one of writing regulations to one of assisting lawmakers to assemble, filter, and focus on the most relevant and pressing information needed to understand a regulatory subject's changing dynamics.[48]

As Hurwitz & Manne note, existing efforts to position some agencies as information-gathering clearinghouses suffer from a number of shortcomings—most notably, that they tend to operate on an *ad hoc* basis, reporting to Congress in response to particular exigencies.[49] The key to developing a "discovery process" for AI regulation would instead require setting up ongoing mechanisms to gather and report on data, as well as directing the process toward "specifications for how information should be used, or what the regulator anticipated to find in the information, prior to its collection."[50]

Embracing regulation as a discovery process means acknowledging the limits of our collective knowledge about AI's potential risks and benefits. This underscores why regulators should prioritize generating and utilizing new information through regulatory experiments, iterative rulemaking, and feedback loops. A more adaptive regulatory framework could respond to new developments and insights in AI technologies, thereby ensuring that regulations remain relevant and effective, without stifling innovation.

Moreover, Hurwitz & Manne highlight the importance of considering regulation as an information-producing activity.[51] In AI regulation, this could involve setting up mechanisms that allow regulators, innovators, and the public to contribute to and benefit from a shared pool of knowledge about AI's impacts. This could include public databases of AI incidents, standardized reporting of AI-system performance, or platforms for sharing best practices in AI safety and ethics.

Static regulatory approaches may fail to capture the evolving landscape of AI applications and their societal implications. Instead, a dynamic, information-centric regulatory strategy that embraces the market as a discovery process could better facilitate beneficial innovations, while identifying and mitigating harms.

---

[47] *Id.* at 33.

[48] *See id.* at 28-29

[49] *Id.* at 37.

[50] *Id.* at 37-38.

[51] *Id.*

V. Conclusion

As the NTIA navigates the complex landscape of AI regulation, it is imperative to adopt a nuanced, forward-looking approach that balances the need to foster innovation with the imperatives of ensuring public safety and ethical integrity. The rapid evolution of AI technologies necessitates a regulatory framework that is both adaptive and principles-based, eschewing static snapshots of the current state of the art in favor of flexible mechanisms that could accommodate the dynamic nature of this field.

Central to this approach is to recognize that the field of AI encompasses a diverse array of technologies, methodologies, and applications, each with its distinct characteristics, capabilities, and implications for society. A one-size-fits-all regulatory model would not only be ill-suited to the task at-hand, but would also risk stifling innovation and hindering the United States' ability to maintain its leadership in the global AI industry. NTIA should focus instead on developing tailored approaches that distinguish between low-risk and high-risk applications, ensuring that regulatory interventions are commensurate with the potential identifiable harms and benefits associated with specific AI use cases.

Moreover, the NTIA must resist the temptation to rely on overly prescriptive definitions of "openness" or to favor particular business models over others. The coexistence of open and closed approaches to AI development is essential to foster a vibrant, competitive ecosystem that drives technological progress and maximizes social benefits. By embracing a flexible regulatory framework that allows for experimentation and iteration, the NTIA can create an environment conducive to innovation while still ensuring that appropriate safeguards are in place to mitigate potential risks.

Ultimately, the success of the U.S. AI industry will depend on the ability of regulators, developers, researchers, and the broader community to collaborate in developing governance frameworks that are both effective and adaptable. By recognizing the importance of open development and diverse business models, the NTIA can play a crucial role in shaping the future of AI in ways that promote innovation, protect public interests, and solidify the United States' position as a global leader in this transformative field.