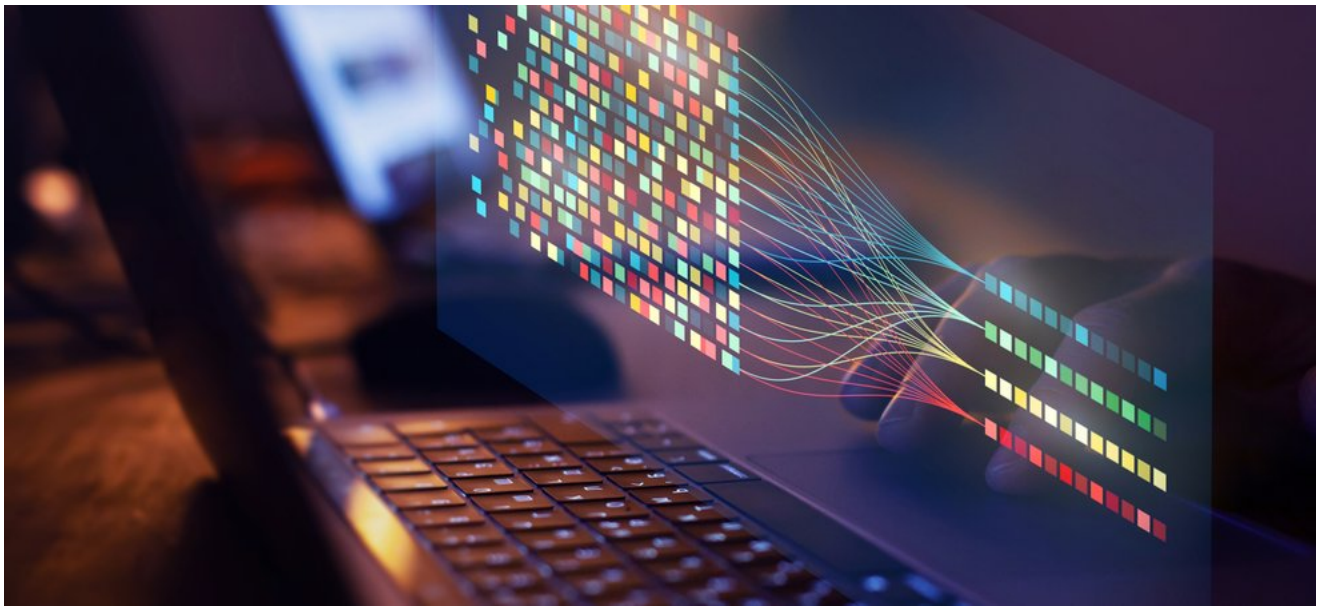
NEW AMERICA



November 2024

Openness in Artificial Intelligence Models

A Key to Ensuring AI Serves Democratic Values and the Public Interest

Prem M. Trivedi & Nat Meysenburg

Open Technology Institute
Last edited on November 19, 2024 at 9:56 a.m. EST

Acknowledgments

We appreciate the insightful comments and edits from Lilian Coral and Emily Tavenner. We are also grateful to Joel Yong for his excellent research assistance.

Editorial disclosure: The views expressed in this report are solely those of the author(s) and do not reflect the views of New America, its staff, fellows, funders, or board of directors.

About the Authors

Prem M. Trivedi is the policy director of New America's Open Technology Institute.

Nat Meysenburg is a technologist at New America's Open Technology Institute.

About New America

We are dedicated to renewing the promise of America by continuing the quest to realize our nation's highest ideals, honestly confronting the challenges caused by rapid technological and social change, and seizing the opportunities those changes create.

About Open Technology Institute

OTI works at the intersection of technology and policy to ensure that every community has equitable access to digital technology and its benefits. We promote universal access to communications technologies that are both open and secure, using a multidisciplinary approach that brings together advocates, researchers, organizers, and innovators.

Contents

Executive Summary

The rapid ascendance of generative artificial intelligence (AI) in today's zeitgeist has spurred policymakers to prioritize governing AI more broadly. In the United States, lawmakers and other stakeholders, including developers and civil society, are considering how AI can better serve democratic institutions, the economy, and consumers. The Biden administration has acted swiftly in issuing the **Blueprint for an AI Bill of Rights** and **Executive Order 14110**, which imposes a detailed array of requirements on federal agencies. In addition to executive action, states have also begun to issue broad and specific laws governing AI, but Congress has yet to take significant action.

Key Attributes of Openness

Society is in the early years of AI's development and even earlier in the approach to governing it. This nascent phase of AI governance presents an opportunity to better understand the concept of "openness" in the context of AI. In this report, we argue that encouraging greater openness in the AI model ecosystem is essential to shaping AI's development in ways that serve democratic values and the public interest.

"Open" and "closed" is not a rigid binary into which AI can be neatly placed. It is more helpful to envision a spectrum of openness. We identify five key attributes of openness for AI models:

1. Open code that can be downloaded, modified, shared, and used by people other than the model's creators;

2. Open licenses that allow third parties to use the model;

3. Transparency about model inputs (data sources, model weights);

4. Transparency about envisioned threats from models and mitigations against undesirable downstream effects (e.g., malicious actors fine-tuning the model to cause clear harms); and

5. Open standards for interconnection and communication among AI models that allow people and companies to switch between models (portability) and for models to interoperate with one another.

Because these attributes of openness encapsulate both technical and non-technical aspects of transparency, they acknowledge that AI models are not

merely software but broader human projects shaped by deliberate governance choices.

Key Benefits of Openness

Many policy discussions about open models narrowly focus on the security risks posed by such models or open weights. A nuanced, empirically grounded discussion of risks is important, but focusing narrowly on risk fails to fully account for the other benefits of openness. Specifically, in the case of AI, promoting the five attributes of openness identified above can create the kind of AI ecosystem that better serves public transparency and democratic accountability, innovation and competition, education and research, and even security. The report discusses the ways in which promoting greater openness in the AI model ecosystem furthers each of these societal benefits.

Recommendations for Promoting Openness

A variety of stakeholders in the United States can play vital roles in creating a more open AI model ecosystem.

- **Policymakers** should continue to build governmental capacity to monitor and mitigate the marginal risks posed by open models. They should also craft legislative and policy requirements that promote transparency about a model's technical elements, as well as its design and governance, and encourage and incentivize developers and companies to build model interoperability. Lastly, they should avoid placing broad restrictions on open models, including through means like export controls, licensing requirements, or broad imposition of liability on developers for downstream harms.

- **Researchers** should comparatively study the organizational structures and practices of teams developing open-source models to identify best practices in development and governance of AI modes. Doing so will enable them to identify resource gaps, prioritize areas of research in the public interest, and articulate use cases private companies or AI labs are unlikely to address because of a lack of commercial interest.

- **AI companies** should embrace openness along multiple axes (code, model weights, model training data, and model interoperability) when developing models. They also should participate and invest in the maintenance of open-source AI projects to help ensure that popular model

projects have adequate resources to find and quickly address security vulnerabilities.

- **Developers** should use best practices in software development that promote both secure code and better insight into a model's structure and training, as well as the decision-making behind those components. In addition, they should explore the design of open protocols and standards for promoting model interoperability.

- **Civil society organizations** should continue to creatively explore the ways in which openness in AI models can further democratic accountability and public-interest objectives and widely communicate these benefits. They should also invest in in-house AI expertise to enable critical insights into how the technology functions and an opportunity to interact with and evaluate open model projects.

Introduction

Since OpenAI's large-language model ChatGPT burst onto the scene in November 2022, society has seen generative artificial intelligence (AI) rapidly shape conceptions of work and life online. Large-language models are an example of generative AI that is built on the recent technical advances in large, multi-purpose "foundation models." But AI encompasses not only the generation of content but also a variety of analytical and predictive tools. AI is a broad umbrella term that people have used for decades "to refer to both a field of study and the machine-based systems that use mathematical models to analyze inputs to complete specific tasks, such as making predictions, recommendations, content, and decisions."[1]

Generative AI's rapid ascendance in the current zeitgeist has spurred policymakers to focus on governing AI more broadly. In the United States, the Biden administration responded swiftly by issuing a Blueprint for an AI Bill of Rights[2] and requirements for federal agencies in Executive Order 14110 on "the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence."[3] As a new presidential administration and Congress prepare to take power, it is critical that—in this time of experimentation—the United States determines how it wants the AI model ecosystem to shape its democracy.

The history of the internet's evolution contains an important example of the consequences of failing to prioritize openness. The internet's early years were deeply "generative," not in the currently popular use of the word as in "generative AI," but rather with the following definition: an environment that fosters wide-ranging creativity and innovation.[4] But due to a confluence of factors[5]—including a lack of prioritizing the promotion of openness—that relatively open era has given way to dominance by a few large companies. This current era of consolidation has eroded the power of internet users, reduced space for new competitors, and constrained the potential for unexpected innovation to emerge from diverse corners. Many of the companies that have dominated this phase of consolidation are now at the center of rapid advances in AI, and society is now at another important juncture in the internet's evolution. How broadly accessible and competitive does the United States want the AI landscape to be? What choices will help ensure that AI best serves democratic institutions and norms globally?

Three broad categories of intervention receive persistent attention in the discourse about how to accomplish this goal: (1) governmental regulation and oversight of AI,[6] (2) developing "public AI" models controlled by non-corporate actors,[7] and (3) ensuring that the AI model ecosystem is sufficiently open in terms of code and other transparency measures. While we point out some of the

intersections between these categories, this report focuses on the ways in which openness can better align AI with serving the public interest.

Many of the policy debates around openness in AI models have focused narrowly on the risks posed by unpredictability in the downstream uses of open-source models. It is indeed important to study the marginal risk posed by open models,[8] but most of the current discourse around risk does not fully account for the benefits that openness can provide. Which lessons about the many benefits of openness in AI models should the United States draw from the long history of open-source software? Which aspects of openness beyond code or model weights should be encouraged? Looking at examples in open-source software can help clarify some of the benefits that openness can bring to AI development. While open AI models and open-source software are not perfectly analogous—in fact, there are important differences between AI and open-source software—many key benefits found in open-source software will transfer to AI.

We do not argue for a reductive one-to-one equation of open-source software development and open societies, but the principles of open software and open models do reflect the ethos of open societies foundational to democracy.[9] The long history of open-source software has demonstrated the importance of openness to several societal benefits that reinforce democratic principles. These benefits include promoting transparency and public accountability, fostering unexpected and iterative innovation, promoting educational and research uses of technology, and bolstering security. All of these benefits are key elements of open societies.

Importantly, the concept of openness in AI models should extend beyond publicly available code or model weights to also encompass the importance of transparency in understanding how technical decisions for models are made and an understanding of who makes them. This broader conceptualization underlies how we use the terms "openness" and "open models."

If policymakers continue to focus disproportionately on the risks of open models, they will help keep the bulk of AI innovation in the hands of a few powerful companies that already dominate social media, cloud, and search capabilities. But this trend is not inevitable. If U.S. policymakers want the benefits of AI to be broadly and equitably distributed and to serve democratic values, then they must consider what kind of AI ecosystem to incentivize and build. An AI ecosystem characterized by open models able to thrive alongside proprietary ones can promote public transparency and accountability, innovation from unexpected corners, new avenues for education and research, and security. To imagine how such an AI ecosystem might look, we first must define the key attributes of an open model.

The Spectrum of Openness

There is no easy binary that opposes "open" and "closed" in the case of artificial intelligence (AI) models. Instead, openness should be viewed as a spectrum.[10] This more flexible understanding of openness fosters productive conversations about both the varied benefits of openness and the marginal risks associated with open models relative to closed models or what is already publicly available online.[11]

"Open-source AI" is a term with a definition that is still in flux and likely to be defined differently by different stakeholders. The Open Source Initiative (OSI), a nonprofit that advocates for the benefits of open source and acts as a multistakeholder standards body that maintains a widely-used definition of open source software as a public benefit, just released its first definition of the term "open source AI."[12] The OSI definition clarifies that to be open, an entire AI system must be considered—both in its "fully functional structure and its discrete structural elements." OSI further notes that "the requirements are the same, whether applied to a **system**, a **model**, **weights and parameters**, or other structural elements."[13]

Efforts such as OSI's to align on a specific definition of open-source AI that is based on certain criteria are helpful, as different actors are currently using the term in different, often self-serving, ways. In a paper last year, David Gray Widder, Sarah West, and Meredith Whittaker described this phenomenon of "openwashing," in which a model developer misleadingly claims the mantle of openness for public relations gains while actually providing access to their model in a way that "should be understood as 'closed.'"[14] They suggest that models fall on "gradients of openness" in which the term "open" can describe models that "offer vastly differing levels of access."

The authors offer three attributes for understanding the openness of models: (1) transparency, (2) reusability, and (3) extensibility. Transparency denotes "the ability to access and vet source code, documentation and data;" reusability is "the ability and licensing needed to allow third parties to reuse source code and/or data;" and extensibility is "the ability to build on top of extant off-the-shelf models, 'tuning' them for one or another specific purpose."[15] These attributes are a useful framework for examining a model's software components.

These attributes also highlight many of the same principles required by OSI's definition. For OSI, an open-source AI system must allow use "for any purpose and without having to ask permission," the ability to study how the system works, the ability to modify it, and to share it "with or without modification." These overlaps suggest a growing agreement around the term "open," particularly the need to include transparency, access, and modification in the definition.

Relevant to the ongoing discussions about how to define open models is the question of access to training data. While some models describe themselves as open and provide code and model weights, they do not provide access to the data used when training the model. A group of scholars recently suggested using the term "open-access AI" in this context, arguing that "'open-source AI' is a misnomer for such models" due to "meaningful differences in access, control, and development."[16] OSI's definition similarly regards access to training data as an essential test in determining whether or not a model is truly open source.

We believe there are at least five key ways in which a model manifests openness, whether it is a large foundation model or a more narrowly tailored one:

1. Open code that can be downloaded, modified, shared, and used by others;
2. Open licenses that allow third parties to use the model;
3. Transparency about model inputs (data sources, model weights[17]);
4. Transparency about envisioned threats from models and ways to mitigate against undesirable downstream effects (e.g., malicious actors fine-tuning the model to cause clear harms); and
5. Open standards for interconnection and communication among AI models that allow people and companies to switch between models (portability) and for models to interoperate with one another.

To illustrate the concept of a spectrum of openness, we offer a simplified breakdown with examples. We have chosen this range of attributes as an exercise in illustratively drawing the line between models, recognizing that the spectrum could consist of many more attributes and reflect greater nuance.

The Spectrum of Openness in AI Models

	Attributes	Model Example(s)
Open	Open code + published model weights + transparency about model inputs (sources, training, methodology, etc.) + transparency about downstream effects	Mistral
	Open code + published model weights + limited/no transparency about model inputs (sources, training data, etc.) + some transparency about downstream effects	Llama
↑	Open code + no published model weights + no transparency about sources and training	Grok
	Closed code + no published model weights + public access + some transparency about model inputs and downstream effects	ChatGPT, Gemini, Claude
Closed	Closed code + proprietary access + no transparency about model inputs or downstream effects	Many companies' internal systems (e.g., for warehouse management and distribution)

NEW AMERICA

The exercise of defining open-source AI, or even placing AI models along such a spectrum, demonstrates that the emerging requirements for an AI model to be considered open are similar to those in earlier free and open-source software projects. To examine—as OSI suggests—the entire system of an AI, we must investigate more attributes than simply the code or the weights. We must think more broadly of models as software projects. The history of open-source software is full of instructive examples of how to (and not to) structure and maintain large software projects, and it is also full of examples of unintended consequences that have shaped tech.

The term "open source" is used throughout this report in a way that includes consideration of all software licensing that meets both the Free Software Foundation's definition of free software[18] and the Open Source Initiative's "Open Source Definition." (OSD).[19] We have chosen to use "open-source" to refer to code as it resonates with the current discourse around open models and not because of a particular preference or recommendation for existing open software licenses. We use the term "open model" throughout this report to echo that discussion but recognize that the lexicon around AI and openness is changing and may ultimately need more terminology—like "open access"—to meaningfully distinguish among model types in the future.

Much of the prevailing discourse around open models focuses on risks and fails to fully account for the significant societal benefits of open models to public transparency and accountability, to unexpected innovation and competition, to education and research, and to security. The following sections explore each of these benefits in further detail.

Benefits of Open-Source AI

Public Transparency and Democratic Accountability

Artificial intelligence (AI) decision-making can be inscrutable in ways that frustrate democratic accountability. The inability to understand how a model reached a decision on any given question presents problems for any process that needs to be auditable, repeatable, and generally subject to external scrutiny.[20] The inability to understand why a model gave the answer it did makes AI unreliable for use in critical applications. By offering access to both the technical components of a model and transparency into the decision-making process of model developers, increased openness in AI offers part of the solution to this inscrutability.

While the technical aspects of this problem, known as explainability, are an important unsolved issue requiring attention, good transparency practices implemented by model developers can help in understanding how certain biases appear in a model and can keep developers accountable for the choices they make in designing and training AI models. Generative AI models trained on data scraped from the public internet often embody the errors and biases in that data, which exacerbates long-standing concerns about algorithmic bias and its discriminatory effects.[21] Increased access to information about data used for training is a vital first step in understanding how bias arises in an AI model.

The risks of AI stem from more than the models—they include the whole systems in which those models are used. As Benjamin Brooks, a fellow at Harvard University's Berkman Klein Center, recently observed in a filing before Australia's Department of Industry, Science, and Resources, "the risk posed by the system will depend on a range of factors, including the intended use-case, specific deployment environment, the extent of human oversight, and the possibility of correction and redress." Similar to the Open Source Initiative, Brooks suggests thinking of AI as a complete system and a model as "a component: a prediction engine."[22] In other words, the details about how a model is set up influence what possible risks it poses. Those risks can be introduced in a variety of ways that extend beyond the model itself, from insecurities in the deployment environment to having no system for correcting the AI when it makes a mistake. An AI system consists of both a model and a software project run by the people who maintain it.

Transparency and accountability comprise one part of addressing AI explainability by offering a framework for evaluating non-technical aspects of a model, such as project management, oversight, and decisions about technical design. Moving beyond what a model is doing technically enables a shift from

only focusing on AI as a new, mysterious technology to helpfully identifying the aspects AI shares with other software. There are many lessons in open source about how to structure and organize a large, globally distributed, technical project in fair and transparent ways. The types of lessons that an AI model project could emulate range from community implementation of best practices for managing code, tracking bugs, codes of conduct, or even building a nonprofit's organizational or legal structure around a single software project.

In addition, open models can be more easily modified and deployed in service of specific public-interest goals than closed models. As Divya Siddarth and Saffron Huang of the Collective Intelligence Project—an initiative that aims to direct technological development toward the collective good—and Audrey Tang, Taiwan's first Minister of Digital Affairs, state, "The open-source community can play a large role here, partnering with democratic innovation organizations to train open models that align with public perspectives."[23] They emphasize that the key features of transparency and open innovation also bolster trust in AI and broaden the base of people who participate in shaping AI's impact on society.[24]

Open code can deliver some transparency benefits on its own, but given the explainability challenge facing all of AI—open code, model weights, and training data should not be misconstrued as a silver bullet for achieving fully explained AI decision-making. Understanding how AI systems arrive at any given answer is still a puzzle that must be solved to build trust in AI. Open models, accompanied by other mechanisms of transparency, make more freely available all of the pieces that AI researchers will need to solve the puzzle of explainability.

But openness alone cannot democratize AI or equitably distribute its benefits throughout society, partly because of the concentration of power in the technology industry[25] and also because privately run AI models cannot perfectly align with specific democratic or public-interest objectives. This realization has increased calls for the development of a "public" AI infrastructure. Like the spectrum between the concepts of open and closed systems, it is also possible to think of a spectrum that spans the concepts of wholly private and wholly public AI models.[26] Proponents of public AI define the term differently but consistently note that government-owned or other non-corporate models can be made democratically accountable in ways that privately owned models cannot.[27] While a public AI infrastructure could be built with closed AI models, in most contexts, the transparency that accompanies open models is aligned with many goals of public AI.

Unexpected Innovation and Competition

In 2008, Jonathan Zittrain, an American professor of internet law and the George Bemis Professor of International Law at Harvard Law School, published *The*

Future of the Internet and How to Stop It. In it, he warned that the internet was losing the "generative" potential of its earlier years and growing increasingly controlled by a few powerful, private gatekeepers. Zittrain wasn't using the term "generative" as many people now do with generative artificial intelligence (AI). He defined generativity as "a system's capacity to produce unanticipated change through unfiltered contributions from broad and varied audiences."[28] He issued a warning that the internet's generativity was under threat:

> "The serendipity of outside tinkering that has marked that generative era gave us the Web, instant messaging, peer-to-peer networking, Skype, Wikipedia—all ideas out of left field. Now it is disappearing, leaving a handful of new gatekeepers in place, with us and them prisoner to their limited business plans.... Even fully grasping how untenable our old models have become, consolidation and lockdown need not be the only alternative. We can stop that future."[29]

Those ideas "out of left field" proved profoundly transformative. The generative internet was worth protecting, but Zittrain's prescient warning became a reality. "Web 2.0," or the social media era of the internet, was dominated by the rise of large social media companies that became intermediaries for much of society's online activity. The resulting environment is less generative. Internet monocultures have spread. A few corporations with access to massive amounts of data enjoy an outsized ability to determine the boundaries of online privacy, shape content creation and consumption, and narrow the range of experiences readily accessible online.

These same companies are at the forefront of today's AI innovation race, and they are poised to extend their consolidation. But an ecosystem in which open models thrive alongside proprietary ones can culturally diversify the technology landscape, promote competition, and spur the kind of unexpected innovation that made the early internet such a powerful force for serving a variety of public-interest objectives.[30]

The past three decades of free and open-source software development have produced examples of lasting innovative impact in large, well-known projects, like the Linux kernel, and in smaller software projects built around the needs of specific communities. Time and again, the insights of people modifying open-source code to fit their own needs and open source's tested value as a cost-effective foundation on which to build have advanced technological developments. The ability to examine and modify code has catalyzed innovative new technologies and changed the landscape of how software vulnerabilities are found and fixed.

Open source's success, and the open standards that make up the internet, provide lessons that can make it easier for AI innovation to be generative in the way

Zittrain meant it. As Leslie Daigle, former chair of the Internet Engineering Task Force Board's oversight committee and internet architecture board, stated in a 2019 white paper: "The more proprietary solutions are built and deployed instead of collaborative open standards-based ones, the less the internet survives as a platform for future innovation."[31] A more recent essay makes a similar point by invoking an ecological imperative "to rewild the internet" in ways that allow spontaneity and broad-based innovation to flourish again.[32] Encouraging the development of more open AI models can play a vital role in such a process. Society cannot foresee the specific innovations that open AI models will create—that is precisely the point—but the lessons of open-source software demonstrate that they will occur and that some will have broad, transformative impacts.

"Society cannot foresee the specific innovations that open AI models will create—that is precisely the point—but the lessons of open-source software demonstrate that they will occur and that some will have broad, transformative impacts."

The Linux kernel serves as an illustrative and concrete example of how open source can have a broad yet unforeseen impact. Linux is everywhere in modern computing and is far more than a platform for tech enthusiasts; billions of non-technical users interact with Linux systems every day. Not only does Google have a long history of running Linux on its servers[33] and having its own customized version of Linux for its developers,[34] but both Google's mobile operating system, Android, and its laptop operating system, ChromeOS, are built using Linux at their core. All Chromebook and Android users are running Linux. Many web hosting providers also use Linux-based systems that run a whole stack of open-source software to provide internet services, including web servers, databases, and remote code processing. These providers run the gamut in size, from small hosting providers to some of the biggest players in hosting, like Amazon Web Services. Open-source web servers make up roughly half of all web servers on the internet.[35] Without Linux and popular open-source projects for tasks like running websites, the internet as we know it may have taken shape differently.

While Google, Amazon, Meta, Microsoft, and Apple certainly all have proprietary "closed source" software, those private tools are often built on top of or interact with open-source software. Recognizing the broader value of an open ecosystem, all of those companies have contributed code, paid labor, and financial or

material support to open-source communities. Even one of the largest historical players in the closed-source software market, Microsoft, has deepened its embrace of open source over the last decade. Notably, the company has simplified the process of running certain open-source software in Windows and purchased GitHub, arguably the world's largest repository of open-source code. There is a shared understanding in tech—from developers of small software projects to large corporate players—that open source has created a solid foundation on which tech innovations, both open and proprietary, are more easily built.

In 1991, when software engineer Linus Torvalds first released the Linux kernel, he could not have predicted that in over three decades, the software would be a cornerstone of the internet or run on billions of smartphones. The year 1991 was still relatively early in the personal computing revolution and the internet was in its commercial infancy; the future was not clear. To the extent that society is in the midst of an AI revolution, that revolution is still in its early days. While experts can speculate about possible paths in AI's development, they should do so with an acknowledgment that unexpected paths are inevitable. By analogy to the 1990s, people are using thirty-pound desktop PCs on dial-up internet and trying to imagine the era of smartphones. There is simply no way to clearly forecast all the surprising ways in which people will use AI over the next 20 years, but based on historical trends, it is very likely that open models will sit at the foundation of some of the biggest advancements in AI.

Importantly, open models built for all sizes and purposes—not simply large foundation models dominating discussion today—will spur critical innovation.[36] As smaller, bespoke models emerge and have transformative impact, the generative potential of open-source innovation in the AI context will become increasingly clear. Open models can be designed and modified to address needs in fields as diverse as medicine, cybersecurity, urban planning, and climate change. A few dominant proprietary models can perhaps be adapted to tackle a subset of these problems, but the interests of large corporations may not naturally align with context-specific applications that serve the public interest. A healthy open-source ecosystem is far more conducive to communities' ability to define problems and solutions on their own terms.

As has been the case with open-source software, open AI models can allow people from various means and backgrounds to respond to use cases in their communities that aren't being considered by the private sector and fill the technical gaps they identify. For example, the **OpenCellular project** is an open-source effort aimed at allowing communities that are not currently served by mobile network operators (MNOs) to form their own MNOs by distributing both the software to build them and hardware schematics.[37] Open-source models can similarly provide a toolset for people who find uses for AI in places where larger AI companies like **OpenAI**, **Anthropic**, or **Google** may not be looking or incentivized to look.

This kind of innovation requires enough time and space to take shape and scale. While the impact of many innovations—like Linux—that are attributable to open source can easily be identified in hindsight, they were not obvious when created. Given the space to develop and the ability to interact with real-word use cases, open-source AI models can follow a similar trajectory as software and catalyze creative new uses.

We must acknowledge that while open models' innovative potential can contribute to a more competitive AI software ecosystem, they will still exist alongside market concentration elsewhere in the AI supply chain. This is particularly true at the infrastructure level, where there are large players who dominate in hardware and cloud computing. Proponents of open models should understand the resource challenges that smaller players currently face in building and training AI, but research toward less resource-intensive model training[38] combined with open-source software's history of enabling competition will produce a more innovative AI ecosystem.

We should not overstate open AI models' ability to serve as a panacea for stopping the consolidation of AI. But the history of open-source software has shown that openness leads to a more innovative environment and brings competitive benefits to the entire tech ecosystem. This lesson should find an analog in AI. Open-source AI models can draw from the history of open-source software, anticipating that the iterative creativity that has brought hard-to-foresee innovation in open-source software for decades will bring about impactful new uses for AI, particularly because those contributions might come from the most unexpected places.

Educational and Research Purposes

One of the key benefits of open source has been its ability to lower the barriers people face when learning how technologies work. The ability to download, study, and freely modify code has led to a wide availability of open-source programming languages and training materials. This has made possible both new forms of training, like coding bootcamps, as well as new avenues for computer science research. Lowering barriers is not only about cost, but also about the ability to freely use and modify code for both training and research purposes.

Open-source models are vital to democratizing education about artificial intelligence (AI) as a technology and avoiding the concentration of knowledge among a small group of people granted access by a few companies to their closed models. Open models and code empower a wider range of people to access these technologies in ways that allow them to gain a hands-on understanding of how they work. Enabling a wider variety of researchers, students, technologists, and hobbyists—along with companies—to examine, run, and modify AI models in unrestricted ways will lead to greater insights about the technology. Such insights

are possible when people are equipped with access to a technology and can form deep knowledge of that technology based on their experiences using it.

Consolidating AI around a few large players with closed-source models runs the risk of confining most high-level AI skills and expertise within the walls of the large tech companies that build those models. Without open models, there would be fewer opportunities for those who are not employed by an AI company—or in that company's approved training pipeline—to learn about the fundamentals of AI technologies. Open models can create multiple pathways to learning AI.

Furthermore, the ability to examine and modify how model code works will lower barriers to academic research on AI. A wider variety of people researching the technology will produce active exploration about a larger and more diverse set of questions about AI.

Indeed, open models may already be impacting AI research, where researchers have already designed experiments using existing open-source AI models to advance the general understanding in the field. For example, Carnegie Mellon and Apple researchers, in 2024, used **Mistral**'s open-source models to explore ways of creating higher quality training regimens for large-language models using synthetic data.[39] What they found was a method for producing more accurate models, using a smaller training corpus and spending less overall time training. This type of research, which could make model training more efficient (and thus less resource intensive), was only possible because the researchers could conduct experiments using an open model. Because they used open tooling and clearly described their methodologies, their more efficient method for producing more accurate models can be replicated, expanded, and used collectively to strengthen AI models for everyone.

In a seminal 1999 article, Lawrence Lessig, the Roy L. Furman Professor of Law and Leadership at Harvard Law School, described the "Open Source Software Moment" taking shape at the time.[40] Laying out the ideals of open source, he observed that "putting into the commons one's work product—of giving away what one makes" might seem "alien to our tradition" but actually functioned much like science where "progress [is] made and given to the next generation."[41] Open models can create multiple pathways to learning AI, and they become a critical part of ensuring AI knowledge is more accessible to people who want to learn about it, whether they are hobbyists or professional researchers. They also make it easier for all of those people to share what they learn with others.

Mitigated Security Risks

Commentary about security in artificial intelligence (AI) often simplistically equates greater model openness with greater risk. Some commentary goes further, claiming that open-source AI is "uniquely dangerous"[42] when compared

to closed models like **ChatGPT**. All AI models carry security risks, but imprecise claims about open AI models miss the greater nuance required when discussing the security risks of AI and how those risks differ between open and closed models. As the National Telecommunications and Information Administration (NTIA) noted in a report required by Executive Order 14110 on "the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence," it is more appropriate to analyze the marginal risks that open models present when compared to closed models and information that is publicly available online.[43]

There is a broad list of harms that could come from AI, including widespread disinformation, but researchers have concluded there is much less evidence about the marginal risks posed by open models.[44] Academic researchers have highlighted certain areas, like the computer generation of child sexual abuse material,[45] where there is evidence that open models pose concerning extra risks. We do not suggest that developers, policymakers, or researchers should take a cavalier approach to such risks but instead propose that they should rigorously study, monitor, and mitigate against marginal risks that do emerge in open models.

The NTIA's report espouses this view, concluding that "the government should **not** restrict the wide availability of model weights for dual-use foundation models at this time."[46] Instead, the NTIA calls for building governmental capacity for monitoring risks and better understanding the benefits of open models. The NTIA report highlights open models' ability to benefit security, including in cyber deterrence and defense, advancing safety research and identifying vulnerabilities, and promoting transparency and accountability through third-party auditing.[47]

AI comes with security concerns, and a responsible approach to building and maintaining any AI model requires vigilance in monitoring potential security vulnerabilities. Security concerns germane to AI fall into three broad categories. The first consists of security concerns that resemble the vulnerabilities that have long been present in other kinds of software development. The second consists of concerns about the downstream uses of AI, which concern how the technology is used rather than how it is developed. The third consists of AI models that could be used to fuel novel cyberattacks.

All code is at risk of having vulnerabilities, regardless of whether developers use an open or closed model for licensing it, and this is likely to hold true even in AI. As researchers at the Wilson Center have noted, "Vulnerabilities can come from dependency management (what, how, and which software packages are pulled into a new software project) to bad-faith actors (people that intentionally break into systems, or contributors intentionally changing the software to be exploitable) whether the software is developed internally or in the open."[48]

An additional source of vulnerability that could be added to this list is simple human error. For example, in some coding languages, an error as small as forgetting to put quotes around a single variable (e.g., $FOO versus "$FOO") could create a vulnerability in the code. The more complex a project's codebase becomes, the more likely that human error will appear somewhere in that project's code or even in one of its dependencies (and this risk increases with every new dependency).

Software vulnerabilities can occur whether or not the source code is open or closed. With AI, as with all software, there is a strong chance that some of the code running it will have errors that introduce vulnerabilities, regardless of the AI model's licensing. However, discourse around open-source models must also account for the ways in which they will confer key security benefits long recognized in open-source software.

Actors from the U.S. private and public sectors have long recognized the ways in which open-source software is central to security. Microsoft's 2023 *Digital Defense Report* emphasizes open-source software's benefits to cybersecurity, noting that the nature of open-source collaboration is key to mapping the threat landscape and scaling responses to those threats.[49] Notably, open-source software has long been critical to the Department of Defense, and military modernization efforts call for broadening the use of open-source software and properly securing it.[50] The Cybersecurity and Infrastructure Security Agency has also highlighted open source's relationship to strong security in the context of open-source AI models, that "the general consensus among the security community is that the benefits of open sourcing... outweigh the harms that might be leveraged by adversaries."[51]

The ability to modify codebases and run AI models independently of their creators is something unique to open models. While some closed models may allow "red teaming" and have a method for bug reporting, security testers may still face constraints on what they can inspect about the model or restrictions on the types of attack they can simulate and when they can simulate them.

By contrast, running code on machinery that is not owned by a model's writer allows security researchers to use a number of security analysis techniques and methods (such as simulating brute force attacks) that might not be available or allowable when evaluating a publicly available closed model like **ChatGPT** or **Gemini**. Such independent security research could uncover ways to manipulate model outputs or alter how an AI system makes choices, which would then make future versions of those models more impervious to such attacks.

In addition to allowing researchers to improve models, the ability to examine an AI model by completely accessing its model weights, training data, and code also allows for formal security vetting by independent third parties. Third parties, whether governmental or private, bring their own goals and lend their reputation to such reviews, which can increase public trust in a model.

Bugs and vulnerabilities will exist in AI code, as they do in all code, but given the strong track record of security in open source, there is a whole class of risks that aren't substantially new or different—such as the environment where the model is run—and there are many well-established methods for addressing them.

The risks that are novel with AI generally manifest "downstream" from the technology itself. That is to say, the harm happens as a result of the application of AI rather than the AI itself. There is much well-founded concern about the use of AI in supercharging disinformation campaigns, but thus far, there is no clear evidence of differential impact based on whether the AI used in such campaigns are open or closed models. Indeed, Microsoft's own research team notes that large-language models like **ChatGPT**, which runs on a closed foundation model, have been weaponized by adversary governments, including China.[52]

There are legitimate security concerns when it comes to AI. These concerns deserve attention and action that addresses them. However, to ensure that a narrow security lens isn't used to hinder the development of an open AI ecosystem, experts must precisely identify where open models present risks beyond publicly available closed models or information that is publicly available online.

Recommendations and Areas for Further Study

There is already a trend toward consolidation of artificial intelligence (AI) in the hands of the largest players currently in tech. The more concentrated the AI industry becomes, the fewer opportunities will exist for learning about how the technology works. There will also be fewer avenues for insight into why a model works the way it does. Models that are more open can help combat the negative effects of an AI market dominated by closed-source models.

Open-source models can build on the open-source principles that have brought so much innovation in tech. Thankfully, this is widely understood by many influential players in tech. Encouragement for open AI models has come from sources as varied as the Cybersecurity and Infrastructure Security Agency, which warns Americans to apply the lessons of open source to open models,[53] and Mark Zuckerberg, who noted that "the bottom line is that open source AI represents the world's best shot at harnessing this technology to create the greatest economic opportunity and security for everyone."[54]

Diverse stakeholders can play vital roles in ensuring that open models thrive in ways that further transparency and accountability, spark innovation, democratize technical education and research, and bolster the security of AI models and systems. They should appreciate the importance of a healthy open AI model ecosystem to transparency and democratic accountability, innovation, competition, and community-driven applications. What follows are specific steps that U.S. policymakers, researchers, AI companies, developers, and civil society organizations can take to promote those objectives.

Policymakers

Legislators and policymakers should incentivize the development of models that maximize openness along its multiple axes and establish rules that require further transparency and accountability of all models, wherever they fall on the spectrum of openness. Taken together with a serious study of marginal risk, the implementation of laws, policies, technical standards, and meaningful transparency norms can produce public accountability and a good governance race to the top.

- **Continue to build governmental capacity to monitor and mitigate the marginal risks posed by open models.**[55]

- **Craft legislative and policy requirements that promote transparency about model design and governance.** These should include access to training data and other inputs that shape model

decision-making, as well as information about how developers and deployers address the risks of model misuse.

- **Encourage and incentivize developers and companies to build model interoperability that promotes standard communication protocols among models.** Doing so promotes collaborative research and the free flow of data and enables people to port data among training models.

- **Avoid placing broad restrictions on open models,** including through means like export controls, licensing requirements, or broad imposition of liability on developers for downstream harms.[56]

Researchers

Research communities are essential to examining and monitoring the openness and health of model systems as well as in identifying AI applications that serve the public interest. Technical researchers also play a role in evaluating the relative efficacy of AI technologies.

- **Engage in comparative studies of the organizational structures and practices of teams developing open-source models**. Which approaches to governance represent best practices? Are they modeled on the structures of other open-source software projects? Are they sustainably resourced?

- **Identify areas of research that the private sector is unlikely to undertake and articulate use cases** private companies or AI labs are unlikely to develop because of a lack of commercial interest. Such research could include improving the efficiency of model training or lowering the hardware thresholds required to run it.

AI Companies

The choices companies developing AI make will have profound impacts on the nature of the AI ecosystem. AI companies that wish to promote greater openness in AI should do the following.

- **Embrace openness along multiple axes when developing models**. The axes include technical elements such as access to code, model weights, and transparency about training data and interoperability among models.

- **Participate and invest in the maintenance of open-source AI projects** to ensure that popular model projects have the resources they need to find and address vulnerabilities in a timely fashion.[57] Addressing resource gaps—perhaps most urgently, compensation for maintainers—builds a foundation on which open-source AI projects can grow.

Developers

Developers have always been critical to the existence and flourishing of open source, and they are essential to building and maintaining a vibrant ecosystem of open models in the age of widespread AI. AI developers can choose how much they share about what goes into the models they are producing and how those decisions are made. They also will be central in developing the protocols and standards that allow for model interoperability, a key feature of promoting competition and consumer choice. Furthermore, developers must understand both the use cases for their AI and the concerns around its overuse and abuse.

- **Use best practices for software development, particularly security practices, that promote both secure code and better transparency** and insight into a model's structure and training.

- **Study and experiment with designing open protocols and standards for moving data between models**, primarily so that models can interoperate more easily. Other uses, such as moving specialized knowledge between models, should also be explored. An appropriate venue for such standards would either need to be found in an existing standard developing organization (SDO), or a new SDO may need to be created along the lines of existing bodies like the Internet Engineering Task Force. Open standards and protocols can help limit anti-competitive lock-in effects and promote research and education.

Civil Society

Civil society plays an indispensable role in monitoring and shaping AI's many uses. Civil society organizations can focus both on identifying the risks that AI systems pose and the benefits of responsible AI use. Civil society can additionally use open models as powerful tools to further their own research, as well as to identify concerns with the technology or its uses.

- **Creatively explore the ways in which openness can further democratic accountability and other public-interest objectives.** They should continue to view AI through a broad lens, taking into account

both risks and possible benefits. Civil society must continue its effective advocacy for the aspects of openness that translate to better accountability and democratic governance, but they also should emphasize the wider range of societal benefits, including innovation, education and research, and security.

- **Invest in in-house AI expertise to enable critical oversight of models**, open or closed, that is based on a better hands-on understanding of how these technologies work.

Conclusion

For the benefits of AI to be broadly and equitably distributed and to serve democratic values, an AI ecosystem must allow open models to thrive alongside proprietary ones. This kind of ecosystem can promote public transparency and accountability, innovation from unexpected corners, new avenues for education and research, and security.

Notes

1 Sarah Forland, "Demystifying AI: A Primer," New America's Open Technology Institute, October 7, 2024, https://www.newamerica.org/oti/blog/demystifying-ai-a-primer/.

2 Biden Administration, "Blueprint for an AI Bill of Rights," White House, https://www.whitehouse.gov/ostp/ai-bill-of-rights/.

3 Office of Science and Technology Policy, *Executive Order 14110, the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence* (White House, October 20, 2023), https://www.whitehouse.gov/briefing-room/presidential-actions/2023/10/30/executive-order-on-the-safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence/.

4 Jonathan L. Zittrain. *The Future of the Internet—and How to Stop It* (Yale University Press and Penguin UK, 2008), https://dash.harvard.edu/bitstream/handle/1/4455262/Zittrain_Future%20of%20the%20Internet.pdf.

5 A non-exhaustive list of these factors should include delays in passing federal privacy legislation, updating pro-competitive legal and regulatory tools, swiftly developing standards for data portability, and purposefully aligning urgent public-interest objectives with needed financial and technical investments.

6 See, e.g., Ami Fields-Meyer and Janet Haven, "Artificial Intelligence, Illiberalism, and the Threat to Democracy," *Foreign Policy*, October 31, 2024, https://foreignpolicy.com/2024/10/31/artificial-intelligence-ai-illiberalism-democracy-civil-rights/.

7 See, e.g., Public AI Network, *Public AI: Infrastructure for the Common Good* (Public AI Network, August 10, 2024), https://publicai.network/whitepaper/PublicAIwhitepaper.pdf; Ganesh Sitaraman and Alex Pascal, "The National Security Case for Public AI," Vanderbilt Policy Accelerator, September 27, 2024, https://cdn.vanderbilt.edu/vu-URL/wp-content/uploads/sites/412/2024/09/27201409/VPA-Paper-National-Security-Case-for-AI.pdf; Nathan Sanders, Bruce Schneier, and Norman Eisen, "How Public AI Can Strengthen Democracy," Brookings, March 4, 2024, https://www.brookings.edu/articles/how-public-ai-can-strengthen-democracy/.

8 See, e.g., *Dual-Use Foundation Models With Widely Available Model Weights* (National Telecommunications and Information Administration, 2024), https://www.ntia.gov/sites/default/files/publications/ntia-ai-open-model-report.pdf.

9 Lawrence Lessig, "Open Code and Open Societies: Values of Internet Governance," *Chicago-Kent Law Review* 74 (February 1999): 1405–1420, https://bit.ly/3UVuyVT.

10 "Several speakers challenged the notion of a binary between 'open' and 'closed' models, pointing toward a spectrum of options regarding the level of access to system components such as datasets, code, model cards, and model weights." Amanda Leal, *Towards Effective Governance of Foundation Models and Generative AI: Takeaways from the Fifth Edition of The Athens Roundtable on AI and the Rule of Law* (Future Society, March 2024), 32, https://thefuturesociety.org/wp-content/uploads/2024/03/Towards-Effective-Governance-of-Foundation-Models-and-Generative-AI.pdf. See also *Dual-Use Foundation Models*, https://www.ntia.gov/sites/default/files/publications/ntia-ai-open-model-report.pdf.

11 "One thing we have already learned is the importance of focusing on the marginal or differential risks and benefits of open weights. For example, we need to measure the risks of open-weight models relative to the risks that already exist today from widely-available information, or from closed models. We have also been encouraged to hear that this is not a binary choice of 'open' vs. 'closed.' Rather there is a broader 'gradient of openness' that we need to consider and that may offer broader options for

policy." Alan Davidson, "National Security and Open Weight Models: Remarks of Alan Davidson," National Telecommunications and Information Administration, March 22, 2024, https://www.ntia.gov/speechtestimony/2024/national-security-and-open-weight-models.

12 "The Open Source AI Definition 1.0," Open Source Initiative, October 28, 2024, https://opensource.org/ai/open-source-ai-definition.

13 "The Open Source AI Definition 1.0," https://opensource.org/ai/open-source-ai-definition.

14 David Gray Widder, Sarah West, and Meredith Whittaker, "Open (For Business): Big Tech, Concentrated Power, and the Political Economy of Open AI," SSRN, August 17, 2023, https://ssrn.com/abstract=4543807.

15 Widder, West, and Whittaker, "Open (For Business)," https://ssrn.com/abstract=4543807.

16 Parth Nobel, Alan Z. Rozenshtein, and Chinmayi Sharma, "Open-Access AI: Lessons From Open-Source Software," Lawfare, October 25, 2024, https://www.lawfaremedia.org/article/open-access-ai--lessons-from-open-source-software.

17 Model weights refer to the numerical value an AI model gives to a piece of information to show the relative strength between it and another piece of information.

18 "What is Free Software? The Free Software Definition," Free Software Foundation, December 27, 2016, https://www.gnu.org/philosophy/free-sw.html.

19 "The Open Source Definition," Open Source Initiative, March 22, 2007, https://opensource.org/osd.

20 See, e.g., Jaden Fiotto-Kaufman, Alexander R. Loftus, et al., "NNsight and NDIF: Democratizing Access to Foundation Model Internals," arXiv, July 18, 2024, https://www.arxiv.org/abs/2407.14561.

21 See, e.g., Spandana Singh, *Charting a Path Forward: Promoting Fairness, Accountability, and Transparency in Algorithmic Content Shaping* (New America's Open Technology Institute, September 9, 2020), https://www.newamerica.org/oti/reports/charting-path-forward/.

22 Benjamin Brooks, "Consultation on Safe and Responsible AI in Australia," Department of Industry, Science and Resources, October 4, 2024, https://consult.industry.gov.au/ai-mandatory-guardrails/submission/view/308.

23 Divya Siddarth, Saffron Huang, and Audrey Tang, "A Vision of Democratic AI," Digitalist Papers, September 22, 2024, https://www.digitalistpapers.com/essays/a-vision-of-democratic-ai.

24 Siddarth, Huang, and Tang, "A Vision of Democratic AI," https://www.digitalistpapers.com/essays/a-vision-of-democratic-ai.

25 "We find that even though there are a handful of meaningfully transparent, reusable, and extensible AI systems, these and all other 'open' AI exists within a deeply concentrated tech company landscape. With scant exceptions that prove the rule, only a few large tech corporations can create and deploy large AI systems at scale…Given the immense importance of scale to the current trajectory of artificial intelligence, this means 'open' AI cannot, alone, meaningfully 'democratize' AI, nor does it pose a significant challenge to the concentration of power in the tech industry." Widder, West, and Whittaker, "Open (For Business)," https://ssrn.com/abstract=4543807.

26 Nik Marda, Jasmine Sun, and Mark Surman, *Public AI: Making AI Work for Everyone, By Everyone* (Mozilla, September 2024), 7, https://assets.mofoprod.net/network/documents/Public_AI_Mozilla.pdf.

27 See, e.g., Marda, Sun, and Surman, *Public AI*, https://assets.mofoprod.net/network/documents/

Public_AI_Mozilla.pdf; Public AI Network, "Public AI," https://publicai.network/whitepaper/PublicAIwhitepaper.pdf; Sitaraman and Pascal, "The National Security Case for Public AI," https://cdn.vanderbilt.edu/vu-URL/wp-content/uploads/sites/412/2024/09/27201409/VPA-Paper-National-Security-Case-for-AI.pdf; Sanders, Schneier, and Eisen, "How Public AI Can Strengthen Democracy," https://www.brookings.edu/articles/how-public-ai-can-strengthen-democracy/.

28 Zittrain, *The Future of the Internet,* https://dash.harvard.edu/bitstream/handle/1/4455262/Zittrain_Future%20of%20the%20Internet.pdf.

29 Zittrain, *The Future of the Internet,* x, https://dash.harvard.edu/bitstream/handle/1/4455262/Zittrain_Future%20of%20the%20Internet.pdf.

30 Rishi Bommasani, Sayash Kapoor, et al., "On the Societal Impact of Open Foundation Models," arXiv, February 27, 2024, https://arxiv.org/pdf/2403.07918.

31 Leslie Daigle, *The Internet Invariants: The Properties Are Constant, Even as the Internet Is Changing* (Thinking Cat, May 16, 2019), 40, https://www.thinkingcat.com/wordpress/wp-content/uploads/2020/08/2019-InvariantsUpdated.pdf.

32 "[The] unpredictability [of…] internet infrastructure makes it generative, worthwhile, and deeply human." Maria Farrell and Robin Berjon, "We Need to Rewild the Internet," *Noēma Magazine*, April 16, 2024, https://www.noemamag.com/we-need-to-rewild-the-internet/.

33 Marc Merlin, "Live Upgrading Thousands of Servers from an Ancient Red Hat Distribution to a 10-Year Newer Debian Based One," presented at the Large Installation System Administration Conference, Washington, DC, November 3–8, 2013, https://www.usenix.org/system/files/conference/lisa13/lisa13-merlin.pdf.

34 Kordian Bruck, Margarita Manterola, and Sven Mueller, "How Google Got to Rolling Linux Releases for Desktops," *Google Cloud* (blog), July 12, 2022, https://cloud.google.com/blog/topics/developers-practitioners/how-google-got-to-rolling-linux-releases-for-desktops.

35 "October 2024 Web Server Survey," Netcraft, October 31, 2024, https://www.netcraft.com/blog/october-2024-web-server-survey/.

36 James Thomason, "Why Small Language Models Are the Next Big Thing in AI," Bloomberg, April 12, 2024, https://www.bloomberg.com/news/articles/2024-08-08/move-over-llms-small-ai-models-are-the-next-big-thing.

37 "OpenCellular," Telecom Infra Project, https://telecominfraproject.com/opencellular/.

38 Paolo Faraboschi, Ellis Giles, Justin Hotard, Konstanty Owczarek, and Andrew Wheeler, "Reducing the Barriers to Entry for Foundation Model Training," arXiv, October 14, 2024, https://arxiv.org/abs/2404.08811v2.

39 Pratyush Maini, Skyler Seto, He Bai, David Grangier, Yizhe Zhang, and Navdeep Jaitly, "Rephrasing the Web: A Recipe for Compute & Data-Efficient Language Modeling," arXiv, January 20, 2024, https://arxiv.org/abs/2401.16380.

40 Lessig, "Open Code and Open Societies, 104, https://bit.ly/3UVuyVT.

41 Lessig, "Open Code and Open Societies," 1411–1412, https://bit.ly/3UVuyVT.

42 David Evan Harris, "Open-Source AI Is Uniquely Dangerous," *Spectrum*, January 12, 2024, https://spectrum.ieee.org/open-source-ai-2666932122.

43 *Dual-Use Foundation Models,* 10, https://www.ntia.gov/sites/default/files/publications/ntia-ai-open-model-report.pdf.

44 Bommasani, Kapoor, et al., "On the Societal Impact of Open Foundation Models," 6, https://arxiv.org/pdf/2403.07918.

45 David Thiel et al., *Generative ML and CSAM: Implications and Mitigations* (Stanford Internet Observatory, June 24, 2023), 7–8, https://stacks.stanford.edu/file/druid:jv206yg3793/20230624-sio-cg-csam-report.pdf.

46 *Dual-Use Foundation Models*, 36, https://www.ntia.gov/sites/default/files/publications/ntia-ai-open-model-report.pdf.

47 *Dual-Use Foundation Models*, 17, https://www.ntia.gov/sites/default/files/publications/ntia-ai-open-model-report.pdf.

48 Ashley Schuett, Alison Parker, and Alex Long, "Open Source Software and Cybersecurity: How Unique Is This Problem?" *CTRL Forward* (blog), Wilson Center, November 10, 2022, https://www.wilsoncenter.org/blog-post/open-source-software-and-cybersecurity-how-unique-problem.

49 Microsoft Threat Intelligence, *Microsoft Digital Defense Report 2023* (Microsoft, October 2023), 116, https://www.microsoft.com/en-us/security/security-insider/microsoft-digital-defense-report-2023.

50 Ben FitzGerald, Jacqueline Parziale, and Peter L. Levin, *Open Source Software and the Department of Defense* (Center for a New American Security, August 30, 2016), https://www.cnas.org/publications/reports/open-source-software-and-the-department-of-defense.

51 Jack Cable and Aeva Black, "With Open Source Artificial Intelligence, Don't Forget the Lessons of Open Source Software," Cybersecurity and Infrastructure Security Agency, July 29, 2024, https://www.cisa.gov/news-events/news/open-source-artificial-intelligence-dont-forget-lessons-open-source-software.

52 Microsoft Threat Intelligence, "Staying Ahead of Threat Actors in the Age of AI," Microsoft Security, February 14, 2024, https://www.microsoft.com/en-us/security/blog/2024/02/14/staying-ahead-of-threat-actors-in-the-age-of-ai/.

53 Jack Cable and Aeva Black, "With Open Source Artificial Intelligence, Don't Forget the Lessons of Open Source Software," https://www.cisa.gov/news-events/news/open-source-artificial-intelligence-dont-forget-lessons-open-source-software.

54 Mark Zuckerberg, "Open Source AI Is the Path Forward," Meta, July 23, 2024, https://about.fb.com/news/2024/07/open-source-ai-is-the-path-forward/.

55 *Dual-Use Foundation Models*, https://www.ntia.gov/sites/default/files/publications/ntia-ai-open-model-report.pdf.

56 For further discussion on how different policy proposals might impact open models, see, e.g., Bommasani, Kapoor, et al., "Considerations for Governing Open Foundation Models," https://www.science.org/doi/10.1126/science.adp1848.

57 Tidelift's 2024 survey of open-source maintainers explains the importance of maintenance to open-source projects and presents responses from over 400 developers. See, e.g., *2024 Tidelift State of the Open Source Maintainer Report* (Tidelift, September 2024), https://tidelift.com/open-source-maintainer-survey-2024.

NEW AMERICA