Issue Brief

Promoting AI Innovation Through Competition

A Guide to Managing Market Power

Author
Jack Corrigan

CSET CENTER for SECURITY and EMERGING TECHNOLOGY

May 2025

Executive Summary

In the years ahead, increasingly advanced artificial intelligence systems are expected to reshape the global economic and national security landscape. Maintaining long-term U.S. leadership in AI will require policymakers to foster a diversified, contestable, and competitive market for commercial AI systems. Competitive markets are a key driver of innovation, incentivizing incumbent firms to continuously improve their offerings and encouraging startups to bring disruptive products to the market. In addition to galvanizing innovation, promoting competition among AI developers would prevent the U.S. AI industry from becoming a digital monoculture and make the country more resilient against economic disruptions, technological surprises, and other national security risks.

As it stands today, however, the commercial AI industry seems poised to become less competitive over time. The organization of the AI supply chain, the economics of AI development, and the industry's prevailing "bigger-is-better" paradigm give incumbent technology companies a clear advantage in the production of AI systems. These firms have both the means and motivation to use their market power to stifle competition in the AI sector. By leveraging their control over computing infrastructure, training data, foundation models, and product distribution channels, incumbents can effectively pick winners and losers in the AI market, potentially preventing disruptive upstarts and their inventions from reaching the market. If left unchecked, this behavior could undermine the long-term innovation capacity and resiliency of the U.S. AI ecosystem.

Though the commercial AI sector is still relatively young, the organization of the market is beginning to solidify. As such, it is appropriate for policymakers to begin pursuing targeted measures to foster competition within the AI ecosystem. Waiting too long to take action risks entrenching current competitive dynamics, allowing large technology incumbents to further extend their market power across the AI industry and potentially stifle domestic innovation. Such inaction could leave the United States at a long-term disadvantage against competitors like China, who have a strong incentive to pursue disruptive innovations that counteract U.S. export controls and other efforts to limit their AI capabilities. This report proposes three broad goals that near-term policy action could target to promote lasting competition and innovation in the AI market, along with specific policy levers for government leaders to consider. These goals include: 1) increasing competition among compute providers; 2) leveling the playing field in the production of AI models and applications; and 3) promoting open distribution channels for AI products.

Table of Contents

Executive Summary 1

Introduction 3

Competition, Innovation, and AI 5

- Promoting Competitive AI Innovation 7
- Competitive Markets and National Security 9

Understanding the AI Market 12

- The AI Supply Chain 12
- Nothing Gold Can Stay: OpenAI and the Economics of AI Development 16
- AI Market Powers 20
  - Chokepoint 1: Computing Resources 21
  - Chokepoint 2: Data Access 23
  - Chokepoint 3: Model Access 24
  - Chokepoint 4: Product Distribution 25
- Key Takeaways 27

Policy Goals for Promoting Competitive AI Innovation 29

- Policy Goal 1: Increase Competition Among Compute Providers 29
- Policy Goal 2: Level the Playing Field in the Production of AI Models and Applications 31
- Policy Goal 3: Maintain Open Distribution Platforms 32
- Trends to Follow 33

Conclusion 35

Authors 36

Acknowledgments 36

Endnotes 37

Introduction

Recent advancements in general-purpose artificial intelligence (AI) systems, especially large language models (LLMs), have captivated leaders across government, industry, and academia. Many tout the technology's ability to improve daily life, transform the economy, and reshape the national security landscape, while others fixate on the technology's risks, such as its potential to entrench social inequalities, empower authoritarian regimes, and instigate global crises. As companies pour billions of dollars into building ever-more powerful AI systems, policymakers are looking for ways to harness the technology's potential benefits while mitigating its purported harms.

To date, AI policy discussions in the United States have focused largely on the production and use of the technology itself: promoting the safe implementation of AI systems, expanding domestic AI infrastructure (semiconductor manufacturing, data centers), and limiting China's ability to scale up similar infrastructure.[1] However, policymakers have paid comparatively less attention to the market conditions in which that technology is being produced.

In the years ahead, it will be critical for policymakers to focus their efforts on promoting AI market competition—namely, the process by which commercial firms vie against one another for dominance in AI. Competition dynamics can significantly impact the rate, diversity, and direction of innovation in a particular market, as well as the ways in which the costs and benefits of those innovations are distributed across society. Indeed, one of the most effective ways for policymakers to influence the trajectory of AI development is by shaping the commercial AI market. As the following section will discuss, competitive AI markets are critical for promoting a diversified, contestable, and competitive innovation ecosystem and maintaining resilience against economic disruptions, technological surprise, and other national security risks.

While many developers have released AI models in recent years, a growing body of research suggests that current technological paradigms and broader forces shaping the AI market give large, incumbent technology companies a distinct advantage in the production and distribution of AI systems, and particularly large pre-trained foundation models.[2] Companies like Alphabet (Google), Amazon, Meta, and Microsoft have both the means and the motivation to use their market power to stifle competition in the AI sector. Many are already taking actions that could make it difficult for new developers—including those producing less resource-intensive AI models and applications—to build and sustain successful independent businesses.[3] Without intervention from policymakers, it is possible the AI market could continue to

consolidate around these firms, potentially hampering long-term AI innovation and exposing the U.S. technology sector to future economic and geopolitical shocks.

Federal policymakers have at their disposal a variety of policy levers that they can use to maintain healthy competition and promote long-term innovation in the AI market. These include traditional competition policy tools like antitrust enforcement, regulation, and intellectual property (IP) policy, as well as other measures like procurement, research funding, and infrastructure development.[4] Under the Biden administration, policymakers began exploring how these tools can be used to cultivate a dynamic, innovative AI industry, and it is critical for the Trump administration to build on this foundation to maintain the United States' global leadership in the technology.[5]

The goal of this brief is to highlight some of the central issues related to competition in the AI market and inform a more holistic approach to AI governance. The first section provides an overview of the relationship between market competition and technological innovation, as well as the role that competitive markets play in maintaining U.S. technological leadership. The second section describes the structure of the AI supply chain and the economic trends shaping the present-day AI market. The third section highlights four major chokepoints in the AI supply chain and ways that incumbent firms can exploit them to advantage themselves and undermine competitors. The fourth section discusses some of the policy tools that can be used to foster competition in the AI industry and spotlights a few technological developments that could affect the competition dynamics in the years ahead.

Competition, Innovation, and AI

Before exploring competition dynamics in the AI market, it is important to first understand why market competition matters from a policy perspective.

One of the primary goals of federal technology and national security policy is to promote innovation.[6] By positioning itself at the forefront of technological progress, the United States can cement its leadership in high-value industries, grow the national economy, and maintain an edge over global competitors. While a variety of factors contribute to innovation, few have been as fundamentally important to American-style ingenuity and progress as market competition.

Market competition—the process by which firms vie against each other for profits and market share—is a cornerstone of capitalism. It is through the process of competition that Adam Smith's "invisible hand" pushes self-interested companies to act in ways that benefit the broader public; in competing for customers, companies are incentivized to lower prices and improve the quality of their products and services.[7] One way that firms try to outmaneuver one another is by investing in innovation, which can include creating new business models, streamlining internal operations, or, most relevant to this brief, developing new technologies. By investing in research and development (R&D) and commercializing new products, firms can gain a leg up on rivals, which lets them expand their slice of the existing market or open up entirely new markets for themselves. Many economists view this process of "creative destruction" as a key driver of technological progress and economic growth.[8]

While competition and innovation are closely intertwined in capitalist economies, the exact relationship between the two is complex; in different situations, market competition may drive firms to invest in innovation or deter them from doing so.

Innovation is an inherently risky process, demanding time, money, and personnel for projects that do not always succeed. To invest in innovation, companies must therefore have both the financial resources to support such efforts and a willingness to absorb the costs of potential failures. Firms in competitive markets may be incentivized to take on these risks, for if they succeed, their new innovation would give them an advantage over their competitors. However, market competition also tends to suppress firms' profit margins, leaving them with less free capital to invest in innovation activities. Additionally, competition may reduce the potential gains from developing new products and processes, as rivals could have an easier time copying those advancements and catching up to the innovator. A widely cited economic study highlighted these tradeoffs, finding that competition initially drives higher rates of

innovation, but excessive competition can dampen this activity.[9] Policymakers must consider this "inverted-U" relationship as they look to promote U.S. leadership in the global AI market.

From a policy perspective, it is also important to understand that not all "innovations" are alike, and different types of companies excel in different aspects of technology development. Researchers have found that larger firms tend to invest more resources in R&D, but their research efforts tend to be less efficient than smaller firms and the innovations they do produce tend to be more "incremental" (e.g., introducing an improved version of a product) than "disruptive" (e.g., introducing the original product).[10] This is not to say that large companies never generate groundbreaking innovations, but rather that established firms tend to focus on advancing existing lines of business, while smaller firms have a greater incentive to innovate in ways that disrupt existing paradigms or serve new sectors in an industry.[11]

Additionally, it is crucial to note that when established firms produce new technologies, without sufficient competitive pressure, they may choose to withhold those inventions from the public. For a historical example consider Bell Labs, the research arm of AT&T during the company's mid-20th century heyday.

Founded in 1925, Bell Labs pioneered inventions that led to the creation of the semiconductor, radar, lasers, solar cells, and other technologies now considered foundational to the modern economy.[12] The lab's capacity for disruptive innovation stemmed at least partly from the market power of its parent company—without AT&T's lucrative long-distance phone monopoly, the company may have pressured Bell Labs to forgo freewheeling basic science in favor of research with more immediate commercial payoffs.[13]

But while AT&T's monopoly helped enable Bell Lab's inventions, it did not necessarily support the commercialization of those discoveries. In fact, it was not until a federal antitrust settlement forced AT&T to start licensing its patents to other companies that many of the lab's discoveries were transformed into usable technologies. This settlement, known as the 1956 Consent Decree, effectively dumped more than 7,800 patents into the market and permitted entrepreneurs to iterate on those ideas without

interference from the corporate behemoth of AT&T.* Beyond helping new companies emerge and develop new products, the agreement also revealed that AT&T had invented a number of technologies—including the answering machine, the fiber optic cable, and the speakerphone—that it chose not to release out of fears that they could undermine the company's business model.[14] The decision is a classic exhibit of the so-called "innovator's dilemma," a situation in which incumbent companies are reluctant to introduce new innovations that could disrupt existing products or lines of business.[15] Researchers have found that competition can compel large incumbents to overcome the innovator's dilemma and more readily pursue disruptive breakthroughs.[16]

Promoting Competitive AI Innovation

This analysis of the economics literature reveals a few general attributes that would characterize a healthy and innovative AI ecosystem, and which could guide the government's future competition policy interventions.

First, promoting a **diversified innovation ecosystem**—one in which many different types of firms are pursuing many different types of products—is important for driving long-term AI progress. Large incumbents and small upstarts approach the technology development process in different ways, and each play a unique role in advancing the state of the art. Large incumbents generally devote more resources to R&D, and their size enables them to more effectively compete with foreign firms in the global economy. However, these firms tend to favor more incremental innovation, and they may not always have an incentive to introduce technologies that disrupt their existing business models. As such, smaller firms and startups are critical players in the innovation ecosystem. Unbeholden to existing business models and driven by a desire

* Specifically, the 1956 Consent Decree required AT&T to license all of its patents on a royalty-free basis. Previously these patents were privately held and, therefore, inaccessible to external companies, entrepreneurs, and inventors. Furthermore, the Consent Decree prohibited AT&T from entering any industry outside of telecommunications, thus preventing the company from using its market power to impede innovators in other sectors. Economists found that opening the company's vaults resulted in a 15% increase in follow-on innovation in fields outside of telecommunications, with small- and medium-sized companies accounting for a significant share of new patenting. Notably, the 1956 Consent Decree has specifically been credited with spurring the creation of the U.S. semiconductor industry, out of which grew Silicon Valley. William Shockley of Shockley Semiconductor and Gordon Teall of Texas Instruments were former Bell Labs employees, for instance, and the Consent Decree allowed them to transform their original work on transistors into new companies. For more, see: Watzinger et al., "How Antitrust Enforcement Can Spur Innovation: Bell Labs and the 1956 Consent Decree"; National Academy of Sciences, "Capitalizing on New Needs and New Opportunities: Government-Industry Partnerships in Biotechnology and Information Technologies" (Washington, DC: The National Academies Press, 2001), https://doi.org/10.17226/10281.

to grow and differentiate themselves from incumbents, these firms are more likely to introduce technologies that disrupt existing markets or create new markets altogether. By constantly injecting new ideas and inventions into the economy, small firms and startups help drive the AI frontier forward, pioneer novel applications, and prevent markets from growing stagnant.

Without this sort of diversification, the U.S. AI sector risks becoming a "digital monoculture," a fragile ecosystem dominated by a few developers that produce nearly identical technologies, and who suffer from similar vulnerabilities and blind spots.[17] The current U.S. market for large pre-trained foundation models already displays some of these characteristics: For years, the industry has been driven by a small cadre of highly-capitalized companies (Amazon, Anthropic, Google, Meta, Microsoft, OpenAI) that have focused on building ever-larger LLMs.[18] An environment in which every U.S. firm pursues progress in the same way not only limits innovation, but it also increases the risk of technological surprise.[19] Policymakers have begun to recognize these dangers in recent months as Chinese AI developers like DeepSeek released products that rival the performance of their U.S. counterparts but are more competitive in terms of price and accessibility.[20] Had the U.S. fostered a more diversified AI ecosystem, such competitors may have emerged domestically rather than across the Pacific.

Second, it is also crucial that policymakers support a **contestable and competitive AI market**—that is, one in which it is relatively easy for firms to enter and exit, and no single firm wields outsized power over the others. If new firms do not have the ability to easily enter the market and build successful lines of business, they will have little incentive to participate in the AI industry. Without that influx of new ideas and talent, the AI market risks growing stagnant and its incumbents complacent; had AT&T faced potential entry from rival telecommunications firms, the company would have likely had more incentive to commercialize inventions like the answering machine or speakerphone in order to maintain its advantage. A diversified AI industry requires a steady stream of entrepreneurs and companies to enter the market, and this inflow will be severely limited if the market is not contestable.

Similarly, fostering competition in the AI market is also crucial for driving long-term innovation. While researchers have found that too much competition can reduce firms' incentive to innovate, insulating companies from competition can have the same effect.[21] Without the threat of rivals stealing market share, established firms have little motivation to pursue disruptive innovation, and even those that do may be deterred from commercializing those discoveries. Economists have long underscored the ways competition can energize markets and compel firms to improve products and diversify their offerings.[22] By promoting a competitive AI ecosystem, policymakers can ensure

commercial developers continue to deliver consistent technological developments and adapt quickly to changes in consumer demand, research breakthroughs, technological paradigms, and geopolitical forces.

The commercial AI sector is still in its infancy, and its products, use cases, and business models are expected to change significantly in the coming years. To ensure this evolving industry continues to deliver the timely, consistent innovation needed to maintain long-term U.S. leadership in AI, policymakers should seek to cultivate a diversified AI ecosystem and promote competitive and contestable AI markets. While the specific policy measures required to achieve these goals will likely change over time, policymakers should generally focus on reducing barriers to entry and ensuring established firms do not exploit their market power in ways that exclude or unfairly disadvantage their competitors.

Competitive Markets and National Security

While a growing community of policymakers recognize the benefits of market competition, the idea that the government should focus on promoting a diversified, contestable, and competitive AI industry remains somewhat controversial in some circles. While these critics make a variety of claims, their arguments typically boil down to the notion that only large companies can marshal the resources and leverage necessary to advance U.S. technology leadership on the global stage, and therefore these firms should be shielded from policy interventions (e.g., antitrust enforcement, regulation) that might upset their current business.[23] Many of the country's most powerful technology companies—including Amazon, Google, Meta, and Microsoft—have invested heavily in lobbying efforts that advance this view.[24]

While this argument may appear self-serving, particularly when promoted by the technology incumbents themselves, it does have some merit. Given their extensive resources, large companies are indeed better equipped to compete in the global economy than small companies. As such, they represent a critical asset for the United States as it competes against China—where firms have close support from the state—and other countries for global AI dominance.[25] Furthermore, as we detail in a later section, the incumbent technology firms control uniquely vast quantities of the data, computational, and financial resources required to produce modern AI systems. These firms are responsible for producing some of the world's most advanced AI models, and

were their dominance to decline, the United States would almost certainly see its leadership in AI and other technologies diminished.*

However, this argument is flawed for a variety of reasons. For one, as previously discussed, insulating AI firms from competition could make the broader industry less dynamic and reduce the speed and scale of AI innovation over the long run. Incumbents could wield their existing market power to keep potential rivals from gaining ground and, free from competitive threats, they would have less incentive to invest in disruptive inventions themselves. Consider how the major U.S. LLM developers had for years generally followed the same resource-intensive approach to AI development, and then were eventually caught off guard by DeepSeek, a Chinese company that, due to U.S. economic controls, had been forced to innovate a more efficient approach to AI development.[26]

The risks of complacent incumbents are well understood within the national security community. Leaders in the Department of Defense (DOD) have long discussed how the consolidation of the defense industrial base has contributed to an overreliance on large primes, which have failed to keep pace with commercial firms in many areas of technology.[27] This problem has driven the department to invest significantly in diversifying the defense industrial base and working with more non-traditional vendors.[28] If the market for particular AI products and services consolidates around a small number of large firms, it could create similar problems for the DOD—particularly as AI tools become increasingly integrated in national security systems.[29]

Beyond harming innovation, concentrated markets create other security risks. For one, overreliance on a small set of companies and products can leave critical infrastructure vulnerable to disruptions. The CrowdStrike outage of July 2024, which disabled IT infrastructure across airports, hospitals, government agencies, and other organizations, put this risk on full display.[30] Leaders in the Department of Defense (DOD) have also noted that concentration in the defense industrial base can restrict access to key components for defense systems and reduce the department's ability to negotiate favorable terms with contractors.[31]

Concentrated markets can also leave the country vulnerable to technological surprise from abroad. Consider the decades-long setbacks that the "Big Three" U.S. automakers faced after the introduction of smaller, more fuel-efficient Japanese vehicles in the

* It is worth noting that even the most aggressive competition policy interventions—corporate breakups—would not eliminate these companies entirely but rather spin off parts of their business into independent or subsidiary entities. While this would impact the original organization's ability to develop AI systems, much of that capacity would still be preserved within the broader economy.

1970s. Had General Motors, Ford, and Chrysler not grown complacent in their market dominance and stopped investing in innovation, some experts argue, they would not have been as vulnerable to this sort of disruption.[32] It would not be hard to imagine a similar future disruption in the AI market, particularly given the way that U.S. firms are doubling down on resource-intensive LLMs while at the same time developers in China are actively exploring alternative AI paradigms.[33]

Finally, shielding incumbent technology firms from all policy interventions (regulation, antitrust enforcement, etc.) in the name of promoting global competitiveness can itself create risks for U.S. national security. Allowing domestic companies to pursue their own interests (i.e., increase profits) often benefits the United States by producing beneficial innovations, expanding the economy, and growing the country's geopolitical influence. However, there are also situations in which corporate interests do not align with national security interests. For example, consider how many prominent AI firms have remained deeply integrated in China even amid growing tensions between Beijing and Washington, and how some of those same firms have pushed back against U.S. policies to hamper China's AI sector. The private sector's pursuit of profit can also sometimes undermine the public good; consider the way social media companies, many of which have now entered the AI industry, knowingly promoted extremist content and downplayed the risks their platforms posed to users' mental health.[34] In situations where corporate interests diverge from national security interests and the public good, it is the responsibility of government leaders to implement policies that align the two. Immunizing private companies from policy interventions for the sake of national security could, paradoxically, create different types of security risks.

Corporate interests and public interests often diverge around issues related to market competition. Private firms have clear incentives to suppress competition and undermine rivals because doing so allows them to increase their profits. However, recognizing that this behavior can, among other things, weaken innovation incentives, discourage entrepreneurship, and reduce economic resilience, policymakers created antitrust laws and other policy levers by which they can reinvigorate competition and improve public outcomes as needed. Incumbent firms will typically object to competition policy interventions because such measures constrain or weaken their market power, but it is critical that policymakers weigh the arguments and interests of those firms against broader national goals. As we saw in the case of AT&T and other antitrust interventions throughout history, penalizing an incumbent firm can clear space for new companies and innovations.[35] Policymakers will likely need to deliver similar short-term disruptions in order to support the long-term growth of a diversified, contestable, and competitive AI industry.

Understanding the AI Market

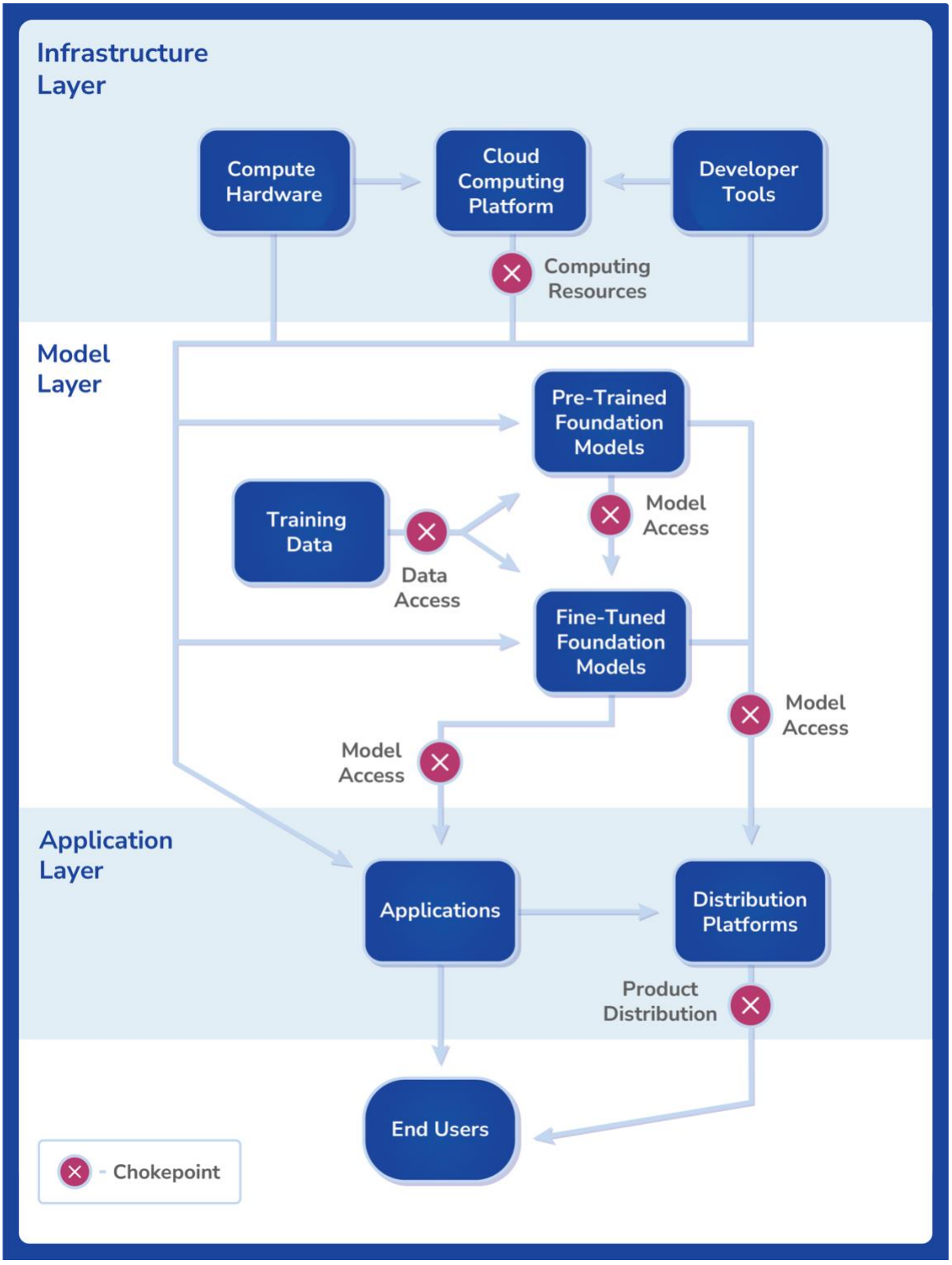
The remainder of this report examines the current AI market based on the broad principles outlined in the prior section: diversification, contestability, and competitiveness. The analysis is intended to explore whether the current dynamics of the industry are aligned with these principles, and how that may change in the future. We begin by examining the AI market itself, describing the structure of the AI supply chain and the economic realities different AI companies face. We then explore how certain incumbent technology companies can exploit their existing assets and market power to suppress competition and advance their own AI products. Finally, we propose three areas where future policy action could create a more open AI market and promote long-term competition and innovation.

The AI Supply Chain

A 2020 CSET report asserted that the many complexities of AI could be boiled down to a single sentence: "Machine learning systems use *computing power* to execute *algorithms* that learn from *data*."[36] While the field of AI has advanced significantly in the ensuing years, this so-called "AI triad"—computing power, algorithms, and data—is still a useful way to frame the essential dynamics of the technology. To understand the ways economic power is distributed and exercised in the AI market, we can examine how the different elements of the triad are brought to bear in the commercial market.

As with other advanced technologies, the AI supply chain is enormously complex. For the purposes of this brief, however, we can break it down into three segments: the **infrastructure layer,** the **model layer,** and the **application layer**. Examining how firms are integrated across each layer can illuminate how and where they can exercise power within the AI market. As shown in Figure 1, each of these layers contain different elements which can be operationalized to produce commercial AI products.

Figure 1. The AI Supply Chain



Source: Author's analysis.

The **infrastructure layer** of the supply chain includes the computing resources that power AI systems, as well as the tools that developers use to train and run those systems. AI developers may access computing resources by either building their own computing clusters or by renting computing infrastructure from a cloud provider. Setting up independent computing clusters for AI development can be an expensive undertaking given the high cost of certain hardware components like advanced semiconductors.[37] As such, many developers rely partly or exclusively on cloud providers, from whom they can rent and scale their computing infrastructure with lower upfront costs.[38] Even so, accessing compute resources is still typically the most expensive part of the AI development process. The modern cloud computing industry is highly concentrated, with just three hyperscalers (the "Big Three") controlling some two-thirds of the global market: Amazon (AWS), Google (GCP), and Microsoft (Azure).* The vast majority of commercial AI developers depend on at least one of these incumbent firms to access computing resources.

There are a variety of software tools that developers can use to build, train, and fine-tune AI systems. Some developer tools are freely available online through platforms like GitHub (owned by Microsoft) and Hugging Face, and others are integrated into proprietary products and services. The Big Three hyperscalers also each offer expansive suites of open-source and proprietary software development tools through their cloud platforms.[39] AI developers generally choose to use the tools provided by their cloud provider, as these solutions are often readily accessible and easy to integrate into the development process.[40]

In the **model layer** of the AI supply chain, developers use infrastructure resources to train algorithms on various datasets with the goal of developing functional AI models. Models may be trained from scratch or developed by altering existing models, and the training process can happen in a number of ways, such as supervised learning, unsupervised learning, or reinforcement learning.[41] The data used to train those models may also come from a variety of sources, including the open internet, pre-existing data repositories, or proprietary digital products (e.g., social media platforms, search engines). Developers may also pay human workers to curate datasets or provide feedback on models.[42]

---

* As of February 2024, the market shares of Amazon Web Services, Microsoft Azure, and Google Cloud Platform were 31%, 24%, and 11%, respectively; "Cloud Market Gets Its Mojo Back; AI Helps Push Q4 Increase in Cloud Spending to New Highs," Synergy Research Group, February 1, 2024, https://www.srgresearch.com/articles/cloud-market-gets-its-mojo-back-q4-increase-in-cloud-spending-reaches-new-highs.

For the purposes of this brief, it is worth differentiating between pre-training and fine-tuning foundation models, because these processes require different amounts of compute and different types of data. Pre-training a foundation model is the initial phase that involves large amounts of diverse data and compute, with the goal of providing the model general knowledge across many domains (i.e., not specifically tailored to any particular application). The second phase is fine-tuning the foundation model so it is specialized in a narrower domain or behaves in a certain way (e.g., tuned to follow instructions or avoid generating toxic outputs). The fine-tuning process often involves training the model on smaller, more curated datasets—an exercise that typically requires far less compute than training a model from scratch.[43]

Once trained, models may be open (i.e., anyone can directly access the underlying model weights), closed (i.e., the public has no access to the underlying model), or gated (i.e., the public has limited access to the underlying model).[44] It is important to note that, because anyone can access and modify open models, developers that rely on such models can skip the resource-intensive pre-training phase and go directly to fine-tuning models for specific tasks or applications. This can, to a degree, reduce the barriers to entry for smaller developers that cannot afford to train models from scratch.

The current U.S. market for pre-trained foundation models is relatively concentrated, with a small number of startups (e.g., Anthropic, Cohere, OpenAI) and large technology companies (e.g., Amazon, Google, Meta, Microsoft) responsible for producing the top-performing models. However, there is a diverse array of fine-tuned models available on the market, offering varying levels of performance and domain expertise.

In the **application layer**, developers package models into software programs that businesses, governments, and other consumers can use. It is important to note that AI applications are not always built by the same developers who trained the underlying model. For example, the paid version of ChatGPT is an AI application that OpenAI built atop its various foundation models, but it is also possible for an independent developer to pay OpenAI for access to those models and use them to build their own applications.

In order to make money, developers also need some way to connect users with their applications. These distribution platforms can take a variety of forms, including online marketplaces and app stores in which AI tools can be proffered directly, as well as other products into which AI tools can be integrated, such as software packages, smartphones, or personal computers. Developers can distribute applications directly to customers, though this strategy provides them significantly less reach than working

through more established channels. As AI systems continue to evolve, the relevance of different distribution platforms may ebb and flow, and new ones may emerge. As we will discuss in a subsequent section, however, today's most prominent distribution platforms—online marketplaces, app stores, enterprise software packages, and personal devices—are generally controlled by incumbent technology firms.

Individual companies may operate across one or more layers of the AI supply chain, and the commercial prospects they face vary widely depending on their level of vertical integration. Google, for example, is integrated across every layer of the AI supply chain: The company operates compute infrastructure (GCP), builds developer tools (e.g., TensorFlow), develops and trains AI models (e.g., Gemini 2.5) with access to proprietary datasets (e.g., Google Search), builds AI applications (e.g., Google AI Assistant), and controls many prominent distribution channels (e.g., Google Play Store, Vertex AI, Android). Because Google's AI supply chain is effectively self-contained, its development costs are relatively low and it is vulnerable to relatively little leverage from other firms.* As we will discuss later, Google can additionally use its compute infrastructure, models, and other assets to gain leverage over other AI firms and generate revenue from these potential competitors. Compare this to a company like Anthropic, which develops and trains AI models (e.g., Claude 3.7 Sonnet) and builds AI applications (e.g., the Claude chatbot) but largely relies on compute infrastructure, distribution channels, and other resources controlled by outside firms. As such, Anthropic is more vulnerable to leverage by other firms and must directly or indirectly pay for access to vital ingredients for AI production.

To understand how the vertical integration of AI companies affects their economic prospects, consider the generative AI industry's breakout startup, OpenAI.

Nothing Gold Can Stay: OpenAI and the Economics of AI Development

In the novel *The Sun Also Rises*, Ernest Hemingway famously wrote that bankruptcy unfolds in two stages: "gradually and then suddenly." Technological change often proceeds along a similar path. Decades of incremental developments, false starts, and

* It is worth noting that Google also designs and uses its own chips (TPUs) to train its AI models, though the company still uses third-party chips from companies like NVIDIA to service its cloud customers. For more information, see: Richard Nieva, "Google Training Gemini on Its Own Chips Reveals Another of Its Advantages," Forbes, December 7, 2023, https://www.forbes.com/sites/richardnieva/2023/12/07/google-deepmind-gemini-tpu/; "NVIDIA and Google Cloud," Google Cloud, accessed September 2024, https://cloud.google.com/nvidia. Chips are a crucial component of the AI supply chain, but an analysis of chokepoints and market power in the semiconductor sector falls beyond the scope of this brief.

disparate discoveries culminate in a breakthrough that launches a period of rapid progress. History offers many examples of these inflection points: the Wright brothers' first flight at Kitty Hawk, the U.S. government's Trinity Test, and the Soviet Union's launch of Sputnik, to name a few. Many people believe we are living through a similar watershed moment in AI. If they are correct, future historians may peg the beginning of this new technological era to November 30, 2022, the day that OpenAI released ChatGPT to the public.[45]

ChatGPT was by no means the first AI tool released into the world—AI systems had been proliferating across smartphones, apps, and websites for years—or even the first LLM, but it captured the public's attention in a way unlike any of the others. Within five days the app had garnered 1 million users, and by January 2023 its user base surpassed 100 million.[46] As the app took off, AI became ingrained in the zeitgeist. Public interest in the technology skyrocketed, investments in AI startups surged, and debates about the promise and perils of AI preoccupied leaders from Silicon Valley to Brussels, Beijing, and Washington.[47]

ChatGPT was significant not just for the popularity and performance of the tool itself, but also the relative obscurity of its inventor. For more than a decade, conversations around AI innovation had largely centered on a handful of industry giants like Amazon, Apple, Google, Meta, and Microsoft. But it was OpenAI, a startup with a relatively meager budget, that instigated the generative AI boom. At first glance, the moment seemed to be a testament to the openness of the AI market, a classic Silicon Valley tale of a nimble startup chasing a crazy idea and disrupting the established players.[48] However, OpenAI's triumph would prove to be short-lived.

While the immense popularity of ChatGPT made OpenAI one of the world's hottest startups, it also saddled the company with major costs. Responding to users' queries required a massive amount of expensive computing resources, and the company had virtually no revenue streams; some estimated OpenAI burned through some $700,000 per day to power a free tool.[49] The company had previously received some $4 billion from various investors, but it needed more funds to stay up and running.[50]

Ultimately the startup found a lifeline in Microsoft. In January 2023, the firms struck a deal in which Microsoft would invest $10 billion worth of cash and cloud computing credits in OpenAI, and in exchange the startup would effectively tether itself to the Seattle tech giant.[51] Specifically, OpenAI agreed to make Microsoft its exclusive provider of cloud services, grant Microsoft an exclusive license to use OpenAI models in its own products, and share 75% of its profits with Microsoft until the company recouped its investment, as well as 49% of its profits thereafter.[52] The deal altered the

landscape of the burgeoning AI industry, propelling Microsoft to the frontier of the model development ecosystem and bringing the sector's breakout startup under the sway of a company founded a decade before OpenAI's CEO, Sam Altman, was born.*

The Microsoft-OpenAI partnership underscored several essential features of the commercial AI industry that are relevant to competition policy.

First, the current structure and economics of the AI supply chain makes **building, deploying, and commercializing AI products a very expensive endeavor.** This is especially true for pre-trained foundation models, which require large amounts of computing power and training data, as well as technical expertise, to bring them to life.[53] The costs of building pre-trained LLMs is expected to continue rising as companies look to increase the scale and performance of their models.[54] Today, many developers are working to build foundation models that require less computational power to develop and deploy.[55] While these "small models" could alter the economics of the AI sector, the modern industry remains dominated by large models, and recent technical developments in scaling "test-time" compute and other AI techniques suggest that the heavy reliance on compute will continue into the future.[56]

For commercial AI companies, the high cost of building AI systems is further exacerbated by the second reality of the market, which is the **ongoing uncertainty about how to monetize AI tools**. While many companies have been able to monetize their AI products through subscriptions and other business models, the path to profitability remains murky; even OpenAI was projected to lose some $5 billion in 2024, two years after it released ChatGPT.[57] In addition to this uncertainty, the recent rollout of generative AI appears to have made little impact on economic productivity, leading some analysts to question whether the technology will be as transformative as initially promised.[58] Recent developments in agentic models and other AI tools could increase the adoption and, thus, the commercial prospects of AI systems.[59] However, today it remains unclear what AI applications will prove most valuable, leaving developers with an uncertain financial future.

This combination of high development costs and uncertain paths to profitability make it exceedingly difficult for commercial AI developers to succeed while maintaining independence from the large vertically integrated technology firms. In the case of OpenAI, the high cost of computing resources threatened to undermine the

* While Microsoft for its part had been conducting AI research and developing AI products on its own for years, partnering with OpenAI gave it a major leg up in the model development ecosystem, as the company could now benefit not only from its own innovation activities but also those at one of the world's top AI labs.

startup's operations—a challenge many AI startups face. Computing infrastructure is perhaps the most widely discussed chokepoint in the AI supply chain, but developers may also face constraints related to data access, model access and distribution platforms. While commercial developers may be able to temporarily cover these costs by raising funds, they are then beholden to investors who often demand near-term returns. Such harsh economic realities can severely constrain the growth options for promising startups, forcing some to seek acquisition by incumbent firms even when they might otherwise prefer to strike out on their own.[60]

Unlike most startups, vertically integrated incumbents control more of the essential inputs for AI development, meaning they can produce AI systems at a much lower cost. Furthermore, given their existing lines of business, such firms can afford to take the calculated risk to lose money on AI in the short run with the expectation that such operations will generate a profit in the medium to long run (similar to AT&T's approach to Bell Labs). This **integration gives large incumbents a distinct advantage over startups in the production of AI systems, particularly large pre-trained foundation models**. Additionally, because these companies often control important AI distribution channels (e.g., online marketplaces, smartphones, popular software packages), they also have a clear advantage over startups in getting their finished AI products into the hands of customers.

Critically, **vertically integrated incumbents can also exploit their role as essential suppliers of AI inputs to gain further dominance over small commercial developers**. Companies like Microsoft, Google, Amazon, and Meta have direct control over the prices, terms, and conditions upon which less-integrated commercial developers can access key components of the AI supply chain, and they can leverage that power to press their advantage, subsume innovation from other companies, and potentially undercut competitors. Microsoft likely could not have negotiated such an advantageous deal with OpenAI had it not operated one of the world's top cloud computing platforms (Azure), a resource the startup very much needed.

The following section will explore some of the ways these vertically integrated incumbents can use their existing assets and market power to tilt the playing field in their favor.[61] While such behavior may be rational from a business perspective, it threatens to stifle competition in the AI market and hamper innovation in the process. By understanding the specific mechanisms through which competition and innovation in the AI market may be undermined, government leaders can better tailor policy interventions to promote a diversified, contestable, and competitive AI market.

AI Market Powers

Over the last 20 years, U.S. Big Tech companies have emerged as arguably the world's most successful economic entities. Their platforms have amassed hundreds of millions, and in some cases billions, of users, and many of their market capitalizations have soared into the trillions. Through this success, the companies have accumulated expansive networks of financial resources, physical infrastructure, and digital assets that position them at the center of the modern economy. Those assets have also turned out to be immensely valuable to the production of AI. As a result, these companies have emerged as some of the top developers of AI systems, producing models and applications that can be directly distributed to users or readily integrated into their other products.

Not only does this integration of assets give Big Tech firms a comparative advantage in the production of AI, but it also positions them as essential suppliers to other developers building AI products. Today it is virtually impossible for an entrepreneur to build and scale a commercial AI business without relying to some degree on resources (i.e., computing infrastructure, software development tools, foundation models, distribution platforms) controlled by a Big Tech company. While such an arrangement is not inherently problematic, it does enable the firms to engage in behavior that reduces competition in the AI market.

The core issue here is that the vertical integration of these firms offers them opportunities to pick winners and losers in the AI market—a market in which they themselves are competing for dominance. Through the selective provision of resources and gatekeeping of various distribution platforms, incumbent tech companies can give certain AI developers and their products an advantage in the marketplace and, conversely, make it difficult for others to compete. Not only do Big Tech firms have the ability to shape the AI marketplace, but because they are also participants in that market, they also have a clear incentive to use their power to give an upper hand to their own products and those of affiliated developers.

It is worth noting that these firms have a long history of engaging in this sort of self-preferencing behavior, as detailed in a variety of media reports, antitrust suits, and government investigations.[62] Such behavior is rational from a business perspective, as it increases firms' market share and allows them to increase profits. However, it may prove harmful from an economic and national security perspective, since stifling competition and excluding new entrants can undermine long-term technological innovation and economic resiliency.

There are a variety of ways that incumbent technology firms can exercise their market power to shape the AI ecosystem, limit competition, and entrench their dominance. For the purposes of this brief, we focus on the ways they can leverage four chokepoints in the AI supply chain to advantage themselves and undermine competitors: computing resources, data access, model access, and product distribution (see Figure 1).

Chokepoint 1: Computing Resources

Perhaps the most widely discussed chokepoint in the AI supply chain is computing infrastructure. Acquiring the hardware and energy needed to train and run large AI models is an expensive endeavor, often one of the costliest parts of building AI systems.[63] As discussed in the prior section, many AI developers find it most economical to access computing resources through cloud service providers.[64] The cloud computing industry is heavily concentrated, with just three hyperscalers—Amazon, Google, and Microsoft—controlling two-thirds of the global market.* This concentration allows these firms to manipulate the economics of AI production in ways that benefit their own businesses and hamper competition.[65]

One clear way that the hyperscalers leverage their control over computing infrastructure to their advantage is by cementing "partnerships" with promising startups, as Microsoft did with OpenAI.[66] In such arrangements, hyperscalers offer startups discounted access to compute infrastructure and other resources in exchange for technology licenses, equity, or other beneficial assets. Such deals often appeal to startups, as they offer them a reprieve from the harsh economics of the AI market and give access to additional support from the incumbent. While many startups are eager to accept such partnerships given the benefits they offer, the problem is that even firms that might otherwise prefer to remain independent may struggle to avoid these arrangements. Investors may pressure the company to accept the financial security that comes from latching themselves to a cloud provider, and startups that refuse such arrangements may struggle to compete with developers who partner with the hyperscalers (and gain access to cheaper compute and other benefits as a result).[67] In effect, these arrangements allow incumbent firms to leverage their existing assets to subsume the innovation and talent of a potential competitor.[68] Industry insiders have suggested these partnerships offer similar benefits to mergers and acquisitions (M&A), but allow incumbents to avoid the regulatory scrutiny associated with traditional M&A deals (though federal regulators have begun investigating these partnerships).[69]

* The Big Three may control a larger collective share of the market among U.S.-based AI developers, who might be reluctant to rely on Chinese-based cloud providers like Alibaba, Tencent, and Huawei.

Aside from startup partnerships, hyperscalers like Amazon, Google, and Microsoft can also leverage their control over computing infrastructure in other ways that make it harder for unaffiliated AI developers to stay in business.[70] Cloud providers may, for instance, charge developers artificially high prices to access their cloud environments.[71] Hyperscalers may also use insights derived from their cloud platform to conduct market research, build copycat products, and engage in predatory pricing.* As we will discuss shortly, incumbent firms can then use their control over distribution platforms to advantage their own AI products over those of their rivals.

Startups would be less vulnerable to this type of conduct if the cloud market was more competitive and they could simply switch from one provider to another. However, the modern cloud computing industry is marked by strong lock-in effects, which means that once a company chooses a cloud provider, it can be expensive to move to a different one.[72] There are technical reasons for these high switching costs; cloud environments are not interoperable, so switching from one to another would require the developer to spend resources retooling their teams and products for the new platform.[73] Cloud providers may also exacerbate these switching costs themselves by charging "egress fees" or locking vendors in restrictive contracts.[74] These lock-in effects not only expose AI startups to the hyperscalers' leverage but also limit their potential customer base; enterprise users—typically large businesses who are themselves locked in to a particular cloud provider—tend to prefer AI tools that can be integrated seamlessly within that cloud environment, making them likely to work only with AI startups that use the same cloud provider. As we will discuss in the following section, some experts have proposed regulations that would reduce these lock-in effects and prevent cloud providers from exploiting their market power to squeeze competitors.[75]

To be sure, not every AI developer is equally vulnerable to the computing infrastructure chokepoint. Different AI systems may require more or less compute, and developers who primarily run inference or work with small models may face lower compute costs than those building large pre-trained foundation models.[76] Firms may also avoid this chokepoint altogether by constructing their own computing infrastructure.[77] However, while compute independence frees developers from cloud

* Amazon has, for instance, been accused of creating knockoffs of highly successful products sold on its e-commerce site and using its control of the platform to advantage the company's own products over those of competitors: Aditya Kalra and Steve Stecklow, "Amazon Copied Products and Rigged Search Results to Promote Its Own Brands, Documents Show," Reuters, October 13, 2021, https://www.reuters.com/investigates/special-report/amazon-india-rigging/; for more information on this practice in an AI context, see: Narechania and Sitaraman, "An Antimonopoly Approach to Governing Artificial Intelligence."

providers' leverage, it is a strategy that only the richest companies can pursue. Meta spent approximately $27 billion on data centers in 2024—an expense it likely could not have afforded had it not already amassed extensive financial and infrastructure assets through its social media businesses.[78]

Chokepoint 2: Data Access

Developing AI models requires vast amounts of training data, though different types of systems demand different amounts and varieties of data. Building large, pre-trained foundation models involves massive troves of diverse data, while fine-tuning models for specific tasks and applications typically requires more narrow, refined datasets. In both cases, incumbent technology firms have a major advantage over startups and small developers in securing this critical resource.

Some training data is freely available to all AI developers through public repositories, but amassing more extensive datasets often requires a significant investment of resources and talent. Scraping and cleaning data from the internet—and fighting the potential legal battles that may ensue—can be an expensive endeavor, and accessing data via third-party providers or crowdsourcing platforms can involve similarly large upfront costs.[79] As smaller AI developers struggle to gain access to these datasets, large incumbents have used their existing capital and market power to amass enormous troves of training data.[80] In some cases, incumbents may outright acquire other companies to gain access to the data generated by their products and services.

Incumbents can also extend this advantage by leveraging data generated by their other lines of business, such as social media platforms, search engines, e-commerce marketplaces, cloud platforms, smart devices, and software suites. Companies collect vast amounts of information on users across these different products and services, and this proprietary data can prove invaluable for training AI systems on individual behaviors, preferences, and patterns. While these data ecosystems are useful for building general-purpose foundation models, they may be even more relevant for fine-tuning models for narrower applications; Google is building tailored AI healthcare tools using data from FitBit, for instance, and Amazon is doing the same using data from One Medical.[81] The incumbents' data advantage can be self-reinforcing: If an incumbent with more training data builds a better AI system, that tool will attract more users, who will in turn generate more data that can be used to improve the AI system. Such feedback loops can confer major first-mover advantages on incumbent tech companies and make it difficult for new rivals to emerge.[82]

Incumbents sometimes share proprietary data with other AI developers, but as with cloud partnerships, these arrangements create opportunities for potentially anti-competitive conduct. Incumbents may share data with certain developers on favorable terms while increasing the barriers to access to other rivals. Consider how Microsoft increased the price of its Bing API shortly after announcing it would integrate OpenAI's models into the search engine—and shortly before DuckDuckGo, a competitor that uses the Bing API, released its own generative search product.[83]

The relative importance of the data access chokepoint could change in the years ahead as the AI industry matures. If the performance of AI systems becomes increasingly correlated to the quality of their training data, as some experts suggest, then incumbents will likely be able to leverage their control over this chokepoint to their advantage.[84] On the other hand, if developers make progress on "small data" techniques or other processes that require less human-generated data to build AI systems, data access could become less of a barrier to AI innovation. Recent developments have shown notable progress on this front: DeepSeek's R1 model, which rivals the performance of its American-made counterparts on certain benchmarks, reportedly employed synthetic data in its training process.[85] While such techniques could help foster a more competitive AI ecosystem, it remains unclear whether synthetic data can be a truly effective substitute for human-labeled data.[86]

Chokepoint 3: Model Access

A third chokepoint that incumbent technology firms can exploit to their advantage is their control over pre-trained foundation models. As discussed above, building large, pre-trained foundation models (e.g., GPT-4o, Gemini 2.5) from scratch can be an immensely expensive endeavor—one that only the most well-resourced firms can currently pursue. As such, many of today's top pre-trained foundation models on the market were developed by U.S. technology incumbents (e.g., Google's Gemini, Meta's Llama, and xAI's Grok) or their affiliated developers (e.g., OpenAI's GPT, Anthropic's Claude).[87] The companies then offer these general-purpose models to downstream developers, who go on to fine-tune them for particular domains or integrate them into new applications or products.

Incumbent firms and their affiliated model providers have full discretion over the terms on which they make their pre-trained foundation models available, which leaves downstream developers—and end users—vulnerable to potentially predatory practices.[88] Model providers may, for instance, arbitrarily raise the price of their products, forcing downstream developers to absorb this additional fee or bear the high costs of switching to a new model.[89] Similarly, model providers may change the terms

of use for their models in ways that harm certain developers, or bundle their models with other products in ways that lock customers into the providers' ecosystem.[90] Furthermore, because model providers also develop AI applications, they may be incentivized to use their control over model access to harm competitors. For example, a model provider may cut off access to a downstream developer that produces a rival AI application or force the developer to share data that helps the provider improve their own rival product.

Different downstream developers may be more or less vulnerable to this chokepoint depending on the "openness" of the foundation model they use. Models that are more open (i.e., available at low or no cost, and/or offer higher degrees of transparency, reusability, and extensibility) are less exploitable than "closed" models, which can be leveraged in the ways described above.[91] Many experts in the AI community and federal government have heralded open models as a way of maintaining a competitive AI market with low barriers to entry.[92] However, in the absence of other competition policy interventions, the mere existence of open models will not, on its own, foster a competitive AI industry. Dominant companies may release open models as a way to attract third-party developers and ensure downstream products are compatible with their product ecosystem, in effect entrenching their power.[93] Indeed, Meta CEO Mark Zuckerberg has said he decided to "open" the company's Llama models for this very reason.[94] While some open models have been released by smaller developers and startups, these models are generally less capable than those of powerful incumbents like Meta, and it is unclear whether these nascent providers will remain financially incentivized to continue building such models as the AI industry matures.

Chokepoint 4: Product Distribution

Perhaps the most under-discussed chokepoint in the AI supply chain is product distribution platforms, or the various avenues through which AI models and applications are marketed, sold, or otherwise made available to users. When building a successful commercial AI business, connecting with users is perhaps just as important as developing innovative products; even the most cutting-edge startup will struggle to stay afloat if it cannot build a customer base. As such, maintaining open distribution platforms for AI products is critical to fostering a competitive and innovative AI ecosystem. Currently, however, some of the most important distribution platforms for AI products (e.g., online marketplaces and app stores, software packages, smartphones, personal computers) are controlled by large incumbent technology companies that can leverage their power to promote the adoption of their own AI systems at the expense of their competitors.

In the modern digital economy, companies rely heavily on third-party platforms to connect with customers. App stores and other online marketplaces act as modern bazaars where vendors can proffer their digital wares, and technologies like smartphones and software suites function as sophisticated multitools into which vendors can integrate their products and gain access to expansive user bases. While the popularity of a particular product depends largely on factors like quality, value, and price, it is also influenced by the platform on which the product is distributed. Items that appear higher in search results tend to draw more attention, for example, just as applications that are set as defaults have an automatic advantage over competing products.[95]

Over the years we have seen many different examples of incumbent technology firms leveraging distribution platforms to advantage certain products over others, with potentially detrimental impacts on competition. A federal judge recently ruled that Google violated antitrust laws when it paid billions of dollars to make itself the default search engine on a variety of devices and web browsers.[96] In 2023, the FTC sued Amazon for, among other things, allegedly manipulating search results to elevate its own products and imposing costly requirements on third-party vendors looking to gain preferential placement on its platform.[97] The U.S. Department of Justice recently sued Apple for reportedly blocking the integration of certain rival applications into its devices, and the agency's antitrust suit against Microsoft in the 1990s focused largely on the ways the company blocked users from accessing rival web browsers.[98]

The current structure of the AI supply chain both enables and incentivizes incumbent platform owners to engage in similar practices with the distribution platforms for AI products. Firms can make AI products developed in-house or by affiliated developers the default applications on their platforms, giving them an automatic advantage over the products of rival developers. We already see examples of this sort of behavior: Google made its Gemini model the default AI assistant on Pixel devices and search engine; Microsoft embedded OpenAI's LLMs in its Microsoft 365 suite; Meta integrated its Llama model into Facebook, Instagram, and WhatsApp.[99] Incumbents may similarly give their affiliated products advantageous placement on other platforms they control, such as app stores, cloud marketplaces, and model hubs. For instance, the homepage for Amazon Bedrock, a generative AI platform for AWS customers, currently lists models from Amazon and Anthropic (in which Amazon is an investor) as its recommended foundation models.[100]

While there is nothing inherently wrong with companies giving preferential placement to their own products on their own platforms, such behavior can make it significantly harder for new developers to gain a foothold in different segments of the market,

especially when those segments are already highly concentrated. Consider that the market for office productivity software—arguably one of the most promising areas for near-term AI deployment—is dominated by Microsoft 365 and Google Workspace.[101] Microsoft and Google have made GPT and Gemini, respectively, the default models for those software suites, potentially excluding other AI developers from a lucrative customer base.[102] Furthermore, effects of self-preferencing may be compounded by the direct and indirect lock-in effects that users experience as AI tools become increasingly attuned to their needs and preferences. Put simply, once a person starts using an AI system, that model will tailor itself to their needs and the user will be less likely to switch to a new product.

While the potential applications of AI are projected to expand in the years ahead, it seems less likely that the channels through which those tools are distributed will change. By leveraging their control over distribution platforms, tech giants can effectively mediate what AI products users are exposed to and, subsequently, which ones will succeed in the digital marketplace. If their past conduct is any indication of future behavior, it seems likely they will use that power to advantage their own AI offerings and suppress those of third-party developers. These exclusionary practices could have a potentially detrimental effect on competition and long-term innovation in the AI market.

Key Takeaways

Today, large incumbent technology companies hold a significant advantage in the production and distribution of AI systems. Having dominated the digital economy for more than a decade, these firms have accumulated deep reserves of the computational infrastructure, data resources, and financial capital required to build AI products. This unmediated access to the raw ingredients for AI production offers incumbents a major competitive edge over smaller companies and new startups, who must pay the incumbents to access at least some of these resources. Crucially, these deep troves of assets also shield incumbents from the harsh economic realities that face those less-resourced developers—namely, high compute costs and uncertain paths to profitability.

Furthermore, the vertical integration of these incumbents across the AI supply chain grants them significant market power, which they can use to entice promising AI developers into acquisition-like corporate "partnerships," promote the distribution and adoption of in-house AI products, and undermine or altogether exclude rival AI developers from the market. The incumbents have a clear financial incentive to use their market power to promote their own AI businesses and suppress those of

competitors, and a variety of media reports, antitrust suits, and government investigations have suggested these firms have willingly engaged in such behavior in the past.

While today a variety of commercial developers participate in the AI market, particularly in the model and application layers, current trends seem poised to push the commercial AI industry towards greater concentration. The AI supply chain is effectively circumscribed by a handful of incumbent technology firms, who control both the resources needed to produce AI systems as well as many of the major channels through which those systems are distributed to users.[103] Should those firms continue to exercise their market power, it will become increasingly difficult for other AI developers to build successful, independent commercial businesses and pursue technologies that are not aligned with the financial interests of the incumbents. The result would likely be a less competitive, less innovative, and less resilient AI ecosystem.

Policy Goals for Promoting Competitive AI Innovation

That said, this concentration of the AI industry is not inevitable. Policymakers have at their disposal a variety of tools that can be used to promote competition in the AI market and prevent incumbent firms from wielding their market power in harmful ways. While some argue that it is still too early for policymakers to intervene in the rapidly developing AI ecosystem, the general organization of the AI supply chain—and the relative market power of its participants—seems to be solidifying. As such, now is precisely the appropriate time for policymakers to pursue targeted measures to promote competition within the AI ecosystem. Waiting too long to take action risks entrenching current competitive dynamics, allowing large technology incumbents to further extend their market power across the AI industry and stifle potentially disruptive innovation. Such inaction could leave the United States at a long-term disadvantage against competitors like China, who have an incentive to pursue disruptive innovations that counteract U.S. export controls and other efforts to limit their AI capabilities.[104]

In this section, we propose three broad goals that near-term policy action could target to promote lasting competition and innovation in the AI market, along with specific policy levers for government leaders to consider. These goals include:

1. Increasing competition among compute providers
2. Leveling the playing field in AI model and application production
3. Maintaining open distribution platforms

We conclude with a discussion of ongoing developments in the AI field that could potentially alter the dynamics of market competition, and which policymakers should continue to monitor in the years ahead.

Policy Goal 1: Increase Competition Among Compute Providers

One area where near-term policy interventions could have a significant impact is the market for computing resources. Today this sector is highly concentrated, with Amazon, Google, and Microsoft controlling two-thirds of the global cloud market, and possibly a larger share of the market for AI-related services. By promoting more competition in this segment of the AI supply chain, policymakers could help drive down the cost of computing resources, making it easier for AI developers to succeed without latching themselves to a cloud provider. Furthermore, subjecting cloud providers to

more competition could promote innovation in cloud services and make it easier for AI developers to better utilize those developments.

To craft effective competition policy for the cloud industry, it is crucial to understand why the sector is so concentrated in the first place. One reason is that the economics of cloud computing tend towards oligopoly: The upfront costs of building cloud infrastructure are very high and, once established, businesses benefit from economies of scale and network effects. These factors have led some researchers to suggest that cloud computing may be a natural monopoly.[105] If this is indeed the case, it may be misguided to pursue policy interventions meant to foster the creation of new private cloud providers; it could take significant time and resources for these new entrants to emerge, and the market could potentially consolidate once again. The introduction of new AI-oriented cloud platforms like CoreWeave and Nvidia's DGX Cloud, as well as the renewed success of Oracle Cloud, in recent years also suggests the industry may not be as impenetrable as some experts have theorized.[106]

A more effective and practical strategy would aim to foster more competition among current cloud providers. Researchers have proposed a variety of policies that could serve this end. For example, policymakers could **mandate interoperability between cloud environments**, which would lower the technical barriers that deter customers from switching providers.[107] Lowering these switching costs would enable AI developers to be more responsive to price changes and improvements in cloud services, which could incentivize cloud providers to reduce prices and invest in innovation. Policymakers could also **prohibit egress fees, restrictive contracting provisions, and other artificial barriers** that cloud providers have themselves erected to lock in customers.[108] Some scholars have also proposed subjecting cloud providers to similar regulations as utility companies, such as prohibitions on price discrimination and other forms of user preferencing.[109]

Other experts have proposed that the federal government could foster healthier competition in the cloud industry by **establishing a public option for cloud services**. This public cloud would offer AI developers an alternative to the existing hyperscalers, giving them more freedom to reject any onerous terms and practices that those private providers might impose. Proponents argue that forcing today's hyperscalers to compete with a more publicly-minded, less profit-oriented entity would compel them to adopt more competitive pricing and limit other predatory practices.[110] Some experts have suggested this government-run cloud could be established through congressional action or the Defense Production Act, and the effort could build on

existing programs like the National Artificial Intelligence Research Resource (NAIRR).* However, unlike the NAIRR, this public cloud would be entirely run by the government.[111] There is precedent for such government-run services. For example, the U.S. Postal Service serves as a public alternative to private postal carriers like FedEx and UPS.[112] Creating a similar public option for cloud services could check the power of incumbent providers and make it easier for new commercial AI businesses to enter the market.

Policy Goal 2: Level the Playing Field in the Production of AI Models and Applications

Another way policymakers can foster long-term competition in the AI market is by leveling the playing field in the production of AI models and applications. While these segments of the AI supply chain today look to be relatively dynamic, prevailing market forces appear to be pushing them toward consolidation. Large incumbent firms have a distinct advantage in the production of AI systems given their existing financial, infrastructure, and digital assets, and they are also already using this market power to subsume promising external developers. Should this behavior continue unchecked, it is likely they will tighten their control over the market for AI models and applications even further. However, there are a variety of ways policymakers can help level the playing field between large incumbents and other AI developers.

One approach could be to **use antitrust enforcement to separate incumbents' cloud operations from the rest of their business**. While there are a number of reasons why incumbent firms like Google, Amazon, and Microsoft excel in the production of AI systems, their ownership of computing infrastructure is one of their most significant advantages. Unlike other developers, which must pay a premium for outsourced compute resources, the Big Three are able to access this infrastructure at marginal cost. Furthermore, these incumbent firms can use the profits from their cloud business to subsidize their AI development efforts. By structurally separating these firms' cloud businesses from the rest of their businesses, policymakers could force large incumbents to compete on a more level playing field with other AI developers. Additionally, policymakers should **continue monitoring traditional M&A and non-traditional corporate "partnerships"** that allow incumbents to subsume independent AI developers, and intervene as necessary.[113]

---

* Dodge, "Creating a Public Cloud Through the Defense Production Act." Unlike the NAIRR, which is being administered by private cloud providers, a public cloud would be operated and owned entirely by the government.

However, antitrust litigation has limitations; cases can take a long time to work through the courts, and by the time a judge reaches their decision, the harm that the defendant inflicted may be irreparable.[114] As such, other policy interventions are necessary to preemptively level the playing field in model and application production.

One such intervention would be to **promote the development of less compute-intensive AI models**. Reducing the resources (computing power, data) required to build high-performing AI systems would lower the barriers to entry for new AI developers and reduce their vulnerability to large incumbents' leverage. Policymakers could help promote this type of innovation by funding research into "small models," data minimization, and other alternative AI paradigms, and requiring the fruits of that research to be publicly shared across the AI community. Shifting the AI industry away from the prevailing "bigger-is-better" paradigm could also reduce the environmental impacts of AI development and increase diversification in the market for AI applications.[115]

Additionally, **supporting the development and proliferation of open models** would ensure developers have viable alternatives to the closed models developed by large incumbents and affiliates, and their associated chokepoints. By supporting the development of open models through federal procurement programs, grants, and other funding mechanisms, policymakers can ensure developers have financial incentives to continue building open models as the AI industry matures. Finally, **enacting data portability requirements** for various digital platforms would help reduce incumbents' data advantage, offer consumers more freedom to switch between AI products, and create a more level playing field on which developers compete for those customers.[116]

Policy Goal 3: Maintain Open Distribution Platforms

Policymakers could also work to maintain open distribution platforms for AI products, which will be key to promoting long-term competition and innovation within the industry. Incumbent firms have already started using their power over distribution platforms like mobile devices, app stores, model hubs, and online marketplaces to make their own products—or those developed by their affiliates—more readily accessible to consumers. In effect, these bundling and self-preferencing practices exclude rival developers from valuable corners of the AI market and make it harder for them to grow their commercial businesses. By **regulating against exclusionary conduct like bundling and self-preferencing**, policymakers can help maintain a more open market in which AI developers compete for customers on the merits of their products rather than their ability to manipulate distribution platforms.[117]

This type of anticompetitive conduct is not unique to the AI sector; policymakers in the United States and abroad have for years criticized large tech companies for leveraging their power over platforms to gain an unfair advantage over competitors. A variety of policy measures have sought to curb this behavior. For example, the European Union's Digital Markets Act prohibits platform "gatekeepers" from, among other things, manipulating search rankings to advantage their own products, and in the United States, congressional lawmakers have introduced a variety of bills to curb this sort of self-preferencing behavior, though none have yet been enacted into law.[118] Applying similar regulations across the AI market could prevent incumbents from unfairly excluding rivals from accessing users. That said, it is unclear what government body would be responsible for implementing and enforcing these regulations, particularly given the United States' lack of a cohesive internet governance framework.

In addition to preemptively regulating against this behavior, federal competition authorities can also **curb exclusionary conduct through antitrust enforcement**. The Justice Department's successful antitrust suit against Google focused in large part on this sort of behavior; specifically, the company's use of contracts that made Google the default search engine on various smartphones and web browsers.[119] However, it is again important to recognize the limitations of antitrust litigation; cases take a long time to process, and because they typically involve a single firm or small group of companies, the impact of remedies can be limited. Maintaining open distribution platforms requires standards and rules that apply systematically across the digital economy, and as such, preemptively regulating such behavior is likely a more scalable long-term solution than antitrust enforcement.

Trends to Follow

This brief has largely focused on the dynamics of competition in the AI market as they stand today. However, the field of AI is advancing rapidly, and future technological and policy developments may change the ways that AI systems are developed and distributed across the economy. These changes may in turn affect the relative importance of different chokepoints in the AI supply chain and, subsequently, the ways in which different firms are able to exercise market power. Tracking these trends and their impacts on different actors in the AI market will be crucial for crafting an effective and durable approach to competition-oriented AI governance. A wide body of research has discussed many of these developments at great length, but here it is worth highlighting a few ongoing trends that could reshape the dynamics of competition in the AI market in the years ahead.

One trend that competition-minded policymakers should follow is the proliferation of so-called "small models." To date, the most powerful AI systems on the market have run on "large models," or algorithms built using significant quantities of computing power and training data. As the energy and hardware costs of computing have soared, however, AI developers have started to experiment with smaller and more efficient algorithms.[120] While today these small models can outmatch their larger counterparts on specific tasks, they generally do not perform as well when measured against a wider array of benchmarks.[121] However, it is possible that this could eventually change, particularly given the strong financial incentives driving firms to develop such systems and advancements in AI orchestration.[122] If small models developed by startups become more prevalent within the AI market, incumbent firms will likely be less able to leverage their control over computing resources to subsume AI developers and undermine competitors, and their comparative advantage in AI production may shrink.

Policymakers should also continue to monitor the proliferation of open models. Developers generally have less control over the distribution and deployment of open models than closed models, which are only available under terms defined by their creators. Government leaders have championed open models as an important way to foster a dynamic and competitive AI ecosystem.[123] Today there is a robust community of developers building and iterating on open models, and the proliferation of such technologies could help maintain low barriers to entry for new AI companies and leave nascent developers less vulnerable to leverage by incumbent developers (whose models are largely more closed).[124] However, open model developers face different monetization challenges than those building closed models, and future AI regulations could directly or indirectly limit the development and proliferation of open models.[125] Should closed models come to dominate the commercial AI market, their developers would likely accumulate more market power.

Finally, policymakers should also monitor the general trajectory of the commercial AI landscape, and specifically the types of models and applications that prove to be most valuable across the economy. The commercial AI industry is still relatively young, and it is still unclear which sectors will adopt AI tools most readily and what the nature of those tools will be. As the demand for different AI systems changes, the relative importance of different AI developers will shift as well. Advancements in certain AI subfields may also change the organization and competition dynamics of the AI market. Maintaining a close eye on the broader AI field and its most prevalent actors, and adapting policies accordingly, will be critical for promoting a competitive and innovative AI ecosystem long into the future.

Conclusion

In the years ahead, increasingly advanced AI systems are expected to reshape the global economic and national security landscape. Maintaining long-term U.S. leadership in this emerging technology will require policymakers to foster a diversified, contestable, and competitive market for commercial AI systems. Competitive markets are a key driver of technological innovation, incentivizing incumbent firms to continuously improve their offerings and empowering startups to bring disruptive new products to the market. Promoting an open and competitive AI market would make the United States less vulnerable to economic disruptions, technological surprise, and other national security risks.

As it stands today, however, the commercial AI industry seems poised to become less competitive over time. The organization of the AI supply chain, the economic realities facing AI developers, and the industry's prevailing "bigger-is-better" technological paradigm give large incumbent technology companies a distinct advantage in the production of AI systems and grants them outsized leverage over their competitors. These incumbents have both the means and motivation to use their existing networks of financial, infrastructural, and digital assets to promote their own AI offerings and undermine the success of their competitors. If left unchecked, we should expect these firms to use their market power to stifle competition in the AI ecosystem, a prospect that could impede long-term innovation and resiliency in the market.

That said, there are a variety of ways that policymakers can intervene in the AI market in ways that address these challenges and promote a more open and competitive AI ecosystem in the years ahead. These include 1) increasing competition among compute providers, potentially through interoperability regulations and public infrastructure investments; 2) leveling the playing field in AI model and application production, potentially through antitrust enforcement, support for open AI models, implementation of data portability standards, and investment in less-resource intensive AI paradigms; and 3) promoting open distribution platforms for AI products by restricting platform owners from engaging in bundling, self-preferencing, and other potentially exclusionary practices. Policymakers should also continue to monitor how future technological and commercial developments in the AI industry impact the dynamics of market competition, and adapt competition policy interventions accordingly.

Authors

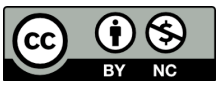
Jack Corrigan is a senior research analyst at Georgetown University's Center for Security and Emerging Technology (CSET).

Acknowledgments

For feedback and assistance, the author would like to thank Sam Bresnick, Owen Daniels, Jessica Ji, Ngor Luong, Kyle Miller, Matthias Oschinski, Emelia Probasco, Christian Schoeberl, Helen Toner, and Igor Mikolic-Torreira.

For helpful conversations, the author would like to thank Rebecca Haw Allensworth, Bill Baer, Cristina Caffara, Pablo Chavez, Adam Conner, Melissa Flagg, Eliana Garces, Campbell Haynes, Tejas Narechania, Howard Shelanski, Ganesh Sitaraman, and Sarah Myers West.

The author is also grateful to Alex Friedland, Jason Ly, and Susan Moynihan for editorial support.



Document Identifier: doi: 10.51593/20230052

Endnotes

[1] Exec. Order No. 14110, 88 FR 75191 (2023); Exec. Order No. 14141, 90 FR 5469 (2025); Hanna Dohmen and Jacob Feldgoise, "A Bigger Yard, A Higher Fence: Understanding BIS's Expanded Controls on Advanced Computing Exports" (Center for Security and Emerging Technology, 2023), https://cset.georgetown.edu/article/bis-2023-update-explainer/; U.S. Department of Commerce, "Biden-Harris Administration Announces CHIPS Incentives Award with Intel to Advance U.S. Leading-Edge Chip Capacity and Create Tens of Thousands of Jobs," November 26, 2024, https://www.commerce.gov/news/press-releases/2024/11/biden-harris-administration-announces-chips-incentives-award-intel; White House, *Framework to Advance AI Governance and Risk Management in National Security* (Washington, DC: The White House, 2024), https://ai.gov/wp-content/uploads/2024/10/NSM-Framework-to-Advance-AI-Governance-and-Risk-Management-in-National-Security.pdf.

[2] Jai Vipra and Sarah Myers West, "Computational Power and AI" (AI Now Institute, September 27, 2023), https://ainowinstitute.org/publication/policy/compute-and-ai; Jon Schmid, Tobias Sytsma, and Anton Shenk, "Evaluating Natural Monopoly Conditions in the AI Foundation Model Market" (RAND, September 12, 2024), https://www.rand.org/pubs/research_reports/RRA3415-1.html; David Gray Widder, Sarah West, and Meredith Whittaker, "Open (For Business): Big Tech, Concentrated Power, and the Political Economy of Open AI," SSRN preprint (August 17, 2023), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4543807; Gaël Varoquaux, Alexandra Sasha Luccioni, and Meredith Whittaker, "Hype, Sustainability, and the Price of the Bigger-is-Better Paradigm in AI," arXiv preprint arXiv:2409.14160 (2024), https://arxiv.org/abs/2409.14160.

[3] Fernando van der Vlist, Anne Helmond, and Fabian Ferrari, "Big AI: Cloud Infrastructure Dependence and the Industrialisation of Artificial Intelligence," *Big Data & Society 11*, no. 1 (2024), https://journals.sagepub.com/doi/10.1177/20539517241232630; Tejas N. Narechania and Ganesh Sitaraman, "An Antimonopoly Approach to Governing Artificial Intelligence," *Yale Law & Policy Review* 43, no. 1 (2024), https://yalelawandpolicy.org/antimonopoly-approach-governing-artificial-intelligence.

[4] Jack Corrigan, Melissa Flagg, and Dewey Murdick, "The Policy Playbook" (Center for Security and Emerging Technology, 2023), https://cset.georgetown.edu/publication/the-policy-playbook/.

[5] Exec. Order No. 14036, 86 FR 36987 (2021); U.S. Federal Trade Commission "FTC, DOJ, and International Enforcers Issue Joint Statement on AI Competition Issues," July 23, 2024, https://www.ftc.gov/news-events/news/press-releases/2024/07/ftc-doj-international-enforcers-issue-joint-statement-ai-competition-issues; Office of Technology, *Partnerships Between Cloud Service Providers and AI Developers* (Washington, DC: U.S. Federal Trade Commission, 2025), https://www.ftc.gov/reports/ftc-staff-report-ai-partnerships-investments-6b-study; David McCabe, "U.S. Clears Way for Antitrust Inquiries of Nvidia, Microsoft and OpenAI," *The New York Times*, June 5, 2024, https://www.nytimes.com/2024/06/05/technology/nvidia-microsoft-openai-antitrust-doj-ftc.html.

[6] Corrigan et al., "The Policy Playbook"; White House, "Memorandum on Advancing the United States' Leadership in Artificial Intelligence; Harnessing Artificial Intelligence to Fulfill National Security Objectives; and Fostering the Safety, Security, and Trustworthiness of Artificial Intelligence," October

24, 2024, https://bidenwhitehouse.archives.gov/briefing-room/presidential-actions/2024/10/24/memorandum-on-advancing-the-united-states-leadership-in-artificial-intelligence-harnessing-artificial-intelligence-to-fulfill-national-security-objectives-and-fostering-the-safety-security/.

[7] Adam Smith, *The Wealth of Nations*.

[8] Joseph Schumpeter, *Capitalism, Socialism, and Democracy* (New York, NY: Harper & Brothers, 1942); Philippe Aghion, Céline Antonin, and Simon Bunel, *The Power of Creative Destruction* (Boston, MA: Harvard University Press, 2023).

[9] Philippe Aghion, Nick Bloom, Richard Blundell et al., "Competition and Innovation: An Inverted-U Relationship," *The Quarterly Journal of Economics 120*, no. 2 (2005): 701–728, https://www.ucl.ac.uk/~uctp39a/ABBGH_QJE_2005.pdf.

[10] John Bound, Clint Cummins, Zvi Griliches et al., "Who Does R&D and Who Patents?," National Bureau of Economic Research, Working Paper Series 908, June 1982, https://www.nber.org/papers/w0908; Wesley M. Cohen, "Chapter 4 — Fifty Years of Empirical Studies of Innovative Activity and Performance," Handbook of the Economics of Innovation, Volume 1, 2010, 129–213, https://doi.org/10.1016/S0169-7218(10)01004-X; Ufuk Akcigit and William R. Kerr, "Growth Through Heterogenous Innovations," *Journal of Political Economy 126*, no. 4 (2018), https://doi.org/10.1086/697901; David Argente, Salome Baslandze, Douglas Hanley, and Sara Moreira, "Patents to Products: Product Innovation and Firm Dynamics," Federal Reserve Bank of Atlanta, Working Paper 2020–4 (2020), https://www.atlantafed.org/-/media/documents/research/publications/wp/2020/04/17/patents-to-products--product-innovation-and-firm-dynamics.pdf; Ufuk Akcigit and Nathan Goldschlag, "Where Have All the Creative Talents Gone? Employment Dynamics of US Inventors," National Bureau of Economic Research, Working Paper Series 31085, March 2023, https://www.nber.org/papers/w31085; Matt Clancy and Arnaud Dyèvre, "The Size of Firms and the Nature of Innovation," *New Things Under The Sun*, June 1, 2023, https://www.newthingsunderthesun.com/pub/d3a1wqfn/release/7; Lingfei Wu, Dashun Wang, and James A. Evans, "Large Teams Have Developed Science and Technology; Small Teams Have Disrupted It," *Nature 566,* 378–382 (2019), https://www.nature.com/articles/s41586-019-0941-9.

[11] Chris Larson, "Disruptive Innovation Theory: What It Is & 4 Key Concepts," Harvard Business School, November 15, 2016, https://online.hbs.edu/blog/post/4-keys-to-understanding-clayton-christensens-theory-of-disruptive-innovation.

[12] Martin Watzinger, Thomas A. Fackler, Markus Nagler, and Monika Schnitzer, "How Antitrust Enforcement Can Spur Innovation: Bell Labs and the 1956 Consent Decree," *American Economic Journal 12*, no. 4 (2020), 328–359, https://pubs.aeaweb.org/doi/pdfplus/10.1257/pol.20190086.

[13] Watzinger et al., "How Antitrust Enforcement Can Spur Innovation: Bell Labs and the 1956 Consent Decree"; Adam D. Thierer, "Unnatural Monopoly: Critical Moments in the Development of the Bell

System Monopoly" (Cato Institute, 1994), https://www.cato.org/sites/cato.org/files/serials/files/cato-journal/1994/11/cj14n2-6.pdf.

[14] Tim Wu, *The Master Switch: The Rise and Fall of Information Empires* (New York, NY: Knopf, 2010).

[15] Clayton Christensen, *The Innovator's Dilemma: When New Technologies Cause Great Firms to Fail* (New York, NY: Harper Business, 2011).

[16] J.P. Eggers and Aseem Kaul, "Motivation and Ability? A Behavioral Perspective on the Pursuit of Radical Invention in Multi-Technology Incumbents," *Academy of Management Journal 61*, no. 1 (2017), https://journals.aom.org/doi/10.5465/amj.2015.1123.

[17] Rory Mir, "CrowdStrike, Antitrust, and Digital Monoculture," Electronic Frontier Foundation, August 1, 2024, https://www.eff.org/deeplinks/2024/07/crowdstrike-antitrust-and-digital-monoculture.

[18] Jack Corrigan, "America Goes All-in on Big AI," *DefenseScoop*, December 17, 2024, https://defensescoop.com/2024/12/17/america-goes-all-in-on-big-ai/.

[19] Aghion et al., "Competition and Innovation: An Inverted-U Relationship."

[20] Dean W. Ball, "Why Manus Matters," *Hyperdimensional*, March 11, 2025, https://www.hyperdimensional.co/p/why-manus-matters; Lina M. Khan, "Stop Worshipping the American Tech Giants," *The New York Times*, February 4, 2025, https://www.nytimes.com/2025/02/04/opinion/deepseek-ai-big-tech.html.

[21] Aghion et al., "Competition and Innovation: An Inverted-U Relationship."

[22] Smith, *The Wealth of Nations*; Schumpeter, *Capitalism, Socialism, and Democracy.*

[23] Robert D. Atkinson, "America Needs Big Tech to Beat China," Information Technology & Innovation Foundation, May 10, 2024, https://itif.org/publications/2024/05/10/america-needs-big-tech-to-beat-big-china/; James Andrew Lewis, "Innovation and Antitrust in Competition with China" (Center for Strategic and International Studies, April 29, 2024), https://www.csis.org/analysis/innovation-and-antitrust-competition-china; Emily Birnbaum, "12 Former Security Officials Who Warned Against Antitrust Crackdown Have Tech Ties," *Politico*, September 22, 2021, https://www.politico.com/news/2021/09/22/former-security-officials-antitrust-tech-ties-513657.

[24] Taylor Giorno, "Big Tech Lobbying Push Helped Block Bipartisan Bills That Aimed to Curb Alleged Anti-Competitive Behavior," *Open Secrets*, December 20, 2022, https://www.opensecrets.org/news/2022/12/big-tech-lobbying-push-helped-block-bipartisan-bills-that-aimed-to-curb-alleged-anti-competitive-behavior/; Will Henshall, "There's an AI Lobbying Frenzy in Washington. Big Tech Is Dominating," *Time*, April 30, 2024, https://time.com/6972134/ai-lobbying-tech-policy-surge/.

[25] Alex Rubin, Alan Omar Loera Martinez, Jake Dow, and Anna Puglisi, "The Huawei Moment" (Center for Security and Emerging Technology, 2021), https://cset.georgetown.edu/publication/the-huawei-

moment/; Edward Zitron, "Deep Impact," *Where's Your Ed At?*, January 29, 2025, https://www.wheresyoured.at/deep-impact/; Ritwik Gupta, "Stop Freaking Out About DeepSeek," January 28, 2025, https://ritwikgupta.me/blog/posts/2025-01-28-stop-freaking-out-about-deepseek/.

[26] Corrigan, "America Goes All-in on Big AI"; Gupta, "Stop Freaking Out About DeepSeek"; Khan, "Stop Worshipping the American Tech Giants."

[27] Marcus Weisgerber, "Lockheed-Sikorsky Deal Stokes Fears About Industry Consolidation," *DefenseOne*, September 30, 2015, https://www.defenseone.com/business/2015/09/sikorsky-lockheed-deal-stokes-fears-about-industry-consolidation/122445/; Mark Thompson, "The Incredibly Shrinking Defense Industry," Project on Government Oversight, August 1, 2019, https://www.pogo.org/analysis/the-incredibly-shrinking-defense-industry; Office of the Under Secretary of Defense for Acquisition and Sustainment, *State of Competition within the Defense Industrial Base* (Arlington, VA: U.S. Department of Defense, 2022), https://media.defense.gov/2022/Feb/15/2002939087/-1/-1/1/STATE-OF-COMPETITION-WITHIN-THE-DEFENSE-INDUSTRIAL-BASE.PDF; Mark Bowden, "The Crumbling Foundation of America's Military," *The Atlantic*, December 17, 2024, https://www.theatlantic.com/politics/archive/2024/12/weapons-production-munitions-shortfall-ukraine-democracy/680867/.

[28] National Defense Science & Technology Strategy Review Task Force, *An Innovation Strategy for the Decisive Decade* (Washington, DC: 2023), https://innovation.defense.gov/Portals/63/DIB_An%20Innovation%20Strategy%20for%20the%20Decisive%20Decade_230717_1.pdf; Melissa Flagg and Jack Corrigan, "Ending Innovation Tourism" (Center for Security and Emerging Technology, 2021), https://cset.georgetown.edu/wp-content/uploads/CSET-Ending-Innovation-Tourism.pdf.

[29] The White House, *Memorandum on Advancing the United States' Leadership in Artificial Intelligence; Harnessing Artificial Intelligence to Fulfill National Security Objectives; and Fostering the Safety, Security, and Trustworthiness of Artificial Intelligence.*

[30] Edward Ongweso Jr, "The Microsoft/CrowdStrike Outage Shows the Danger of Monopolization," *The Guardian*, July 20, 2024, https://www.theguardian.com/technology/article/2024/jul/20/the-microsoftcrowdstrike-outage-shows-the-danger-of-monopolization.

[31] Office of the Under Secretary of Defense for Acquisition and Sustainment, *State of Competition within the Defense Industrial Base,* 27.; Matt Stoller and Lucas Kunce, "America's Monopoly Crisis Hits the Military," *The American Conservative*, June 27, 2019, https://www.theamericanconservative.com/americas-monopoly-crisis-hits-the-military/.

[32] Paul Ingrassia, *Crash Course: The American Automobile Industry's Road from Glory to Disaster* (New York, NY: Random House, 2010).

[33] Corrigan, "America Goes All-in on Big AI"; William Hannas, Huey-Meei Chang, Max Riesenhuber, and Daniel Chou, "China's Cognitive AI Research" (Center for Security and Emerging Technology, 2023), https://cset.georgetown.edu/publication/chinas-cognitive-ai-research/.

[34] Georgia Wells, Jeff Horwitz, and Deepa Seetharaman, "Facebook Knows Instagram Is Toxic for Teen Girls, Company Documents Show," *The Wall Street Journal*, September 14, 2021, https://www.wsj.com/articles/facebook-knows-instagram-is-toxic-for-teen-girls-company-documents-show-11631620739; Karen Hao, "How Facebook Got Addicted to Spreading Misinformation," *MIT Technology Review*, March 11, 2021, https://www.technologyreview.com/2021/03/11/1020600/facebook-responsible-ai-misinformation/.

[35] Khan, "Stop Worshipping the American Tech Giants"; Watzinger et al., "How Antitrust Enforcement Can Spur Innovation: Bell Labs and the 1956 Consent Decree"; Giovanna Massarotto, "Driving Innovation with Antitrust," *ProMarket.org*, April 10, 2024, https://www.promarket.org/2024/04/10/driving-innovation-with-antitrust/.

[36] Ben Buchanan, "The AI Triad and What It Means for National Security Strategy" (Center for Security and Emerging Technology, 2020), https://cset.georgetown.edu/publication/the-ai-triad-and-what-it-means-for-national-security-strategy/.

[37] A single Nvidia A100 can range from $8,000 to $20,000 depending on the model. For more information, see: Margaret Shen, "How Much Is an Nvidia A100?" Modal.com, October 31, 2024, https://modal.com/blog/nvidia-a100-price-article.

[38] Joe Maglitta, "Purpose-Built AI Hardware: Smart Strategies for Scaling Infrastructure," *VentureBeat*, January 16, 2025, https://venturebeat.com/ai/purpose-built-ai-hardware-the-key-to-scalable-ai-infrastructure/; Guido Appenzeller, Matt Bornstein, and Martin Casado, "Navigating the High Cost of AI Compute," *Andreessen Horowitz*, April 27, 2023, https://a16z.com/navigating-the-high-cost-of-ai-compute/; Johan Moreno, "70% Of Generative AI Startups Rely On Google Cloud, AI Capabilities," *Forbes*, July 25, 2023, https://www.forbes.com/sites/johanmoreno/2023/07/25/70-of-generative-ai-startups-rely-on-google-cloud-ai-capabilities-says-alphabet-ceo-sundar-pichai/.

[39] Van der Vlist et al., "Big AI: Cloud Infrastructure Dependence and the Industrialisation of Artificial Intelligence."

[40] Christian Schoeberl and Jack Corrigan, "Funding the AI Cloud—Amazon, Alphabet, and Microsoft's Cloud Computing Investments, Part 1," Center for Security and Emerging Technology (blog), October 30, 2024, https://cset.georgetown.edu/publication/funding-the-ai-cloud-computing-investments-part-1/; Christian Schoeberl and Jack Corrigan, "Funding the AI Cloud—Amazon, Alphabet, and Microsoft's Cloud Computing Investments, Part 2," Center for Security and Emerging Technology (blog), November 13, 2024, https://cset.georgetown.edu/publication/funding-the-ai-cloud-amazon-alphabet-and-microsofts-cloud-computing-investments-part-2/.

[41] Buchanan, "The AI Triad and What It Means for National Security Strategy"; Kyle Miller, "Open Foundation Models: Implications of Contemporary Artificial Intelligence" (Center for Security and Emerging Technology, March 12, 2024), https://cset.georgetown.edu/article/open-foundation-models-implications-of-contemporary-artificial-intelligence/.

[42] Courtney Radsch, "Dismantling AI Data Monopolies Before It's Too Late," *Tech Policy Press*, October 9, 2024, https://www.techpolicy.press/dismantling-ai-data-monopolies-before-its-too-late/; Pablo Villalobos, Anson Ho, Jaime Sevilla et al., "Will We Run Out of Data? Limits of LLM Scaling Based on Human-Generated Data," arXiv preprint arXiv:2211.04325 (2024), https://arxiv.org/pdf/2211.04325; Supantha Mukherjee and Anna Tong, "If your AI seems smarter, it's thanks to smarter human trainers," *Reuters*, September 28, 2024, https://www.reuters.com/technology/artificial-intelligence/if-your-ai-seems-smarter-its-thanks-smarter-human-trainers-2024-09-28/; Niamh Rowe, "Millions of Workers Are Training AI Models for Pennies," *Wired*, October 16, 2023, https://www.wired.com/story/millions-of-workers-are-training-ai-models-for-pennies/.

[43] Edouard d'Archimbaud, "The Challenges, Costs, and Considerations of Building or Fine-Tuning an LLM," *Hackernoon*, September 1, 2023, https://hackernoon.com/the-challenges-costs-and-considerations-of-building-or-fine-tuning-an-llm.

[44] In reality, the "openness" of an AI system exists along a spectrum rather than in discrete categories. For more information, see: Widder et al., "Open (For Business): Big Tech, Concentrated Power, and the Political Economy of Open AI"; Miller, "Open Foundation Models: Implications of Contemporary Artificial Intelligence"; Irene Solaiman, "The Gradient of Generative AI Release: Methods and Considerations," arXiv preprint arXiv:2302.04844 (2023), https://arxiv.org/abs/2302.04844.

[45] "Introducing ChatGPT," *OpenAI*, November 30, 2022, https://openai.com/blog/chatgpt.

[46] Steve Mollman, "ChatGPT Gained 1 Million Users in Under a Week. Here's Why the AI Chatbot Is Primed to Disrupt Search as We Know It," *Forbes*, December 9, 2022, https://finance.yahoo.com/news/chatgpt-gained-1-million-followers-224523258.html; Krystal Hu, "ChatGPT Sets Record for Fastest-Growing User Base — Analyst Note," *Reuters*, February 2, 2023, https://www.reuters.com/technology/chatgpt-sets-record-fastest-growing-user-base-analyst-note-2023-02-01/.

[47] Data Source: Google Trends (https://trends.google.com/trends/explore?date=2023-01-01%202023-12-31&geo=US&q=generative%20AI&hl=en); Leah Hodgson, "VCs Go Vertical in Backing Specialized AI," *PitchBook*, January 29, 2024, https://pitchbook.com/news/articles/vertical-ai-vc; Bruce Schneier and Nathan Sanders, "The A.I. Wars Have Three Factions, and They All Crave Power," *The New York Times*, September 28, 2023, https://www.nytimes.com/2023/09/28/opinion/ai-safety-ethics-effective.html.

[48] Nico Grant and Cade Metz, "A New Chat Bot Is a 'Code Red' for Google's Search Business," *The New York Times*, December 21, 2022, https://www.nytimes.com/2022/12/21/technology/ai-chatgpt-google-search.html.

[49] Dylan Patel and Afzal Ahmed, "Peeling the Onion's Layers—Large Language Models Search Architecture and Cost," *SemiAnalysis*, February 13, 2023, https://www.semianalysis.com/p/peeling-the-onions-layers-large-language.

[50] Data provided by PitchBook Data, Inc.

[51] Microsoft and OpenAI have not disclosed the share of the investment that came in the form of cloud credits: Reed Albergotti, "OpenAI Has Received Just a Fraction of Microsoft's $10 Billion Investment," *Semafor*, November 18, 2023, https://www.semafor.com/article/11/18/2023/openai-has-received-just-a-fraction-of-microsofts-10-billion-investment; Microsoft had previously invested some $3 billion in the company: Cade Metz and Karen Weise, "Microsoft Bets Big on the Creator of ChatGPT in Race to Dominate A.I.," *The New York Times*, January 12, 2023, https://www.nytimes.com/2023/01/12/technology/microsoft-openai-chatgpt.html.

[52] Tom Warren, "Microsoft Extends OpenAI Partnership in a 'Multibillion Dollar Investment'," *The Verge*, January 23, 2023, https://www.theverge.com/2023/1/23/23567448/microsoft-openai-partnership-extension-ai. In June 2024, Microsoft agreed to allow OpenAI to partner with Oracle for additional cloud computing resources: Cade Metz, Mike Isaac, and Erin Griffith, "Microsoft and OpenAI's Close Partnership Shows Signs of Fraying," *The New York Times*, October 17, 2024, https://www.nytimes.com/2024/10/17/technology/microsoft-openai-partnership-deal.html; "OpenAI Selects Oracle Cloud Infrastructure to Extend Microsoft Azure AI Platform," *Oracle*, June 11, 2024, https://www.oracle.com/news/announcement/openai-selects-oracle-cloud-infrastructure-to-extend-microsoft-azure-ai-platform-2024-06-11/.

[53] Companies do not generally publicize the cost of training their models, but researchers estimate that developers have spent tens or even hundreds of millions of dollars training some of today's top models. For more, see: Ben Cottier, Robi Rahman, Loredana Fattorini, Nestor Maslej, Tamay Besiroglu, David Owen, "The Rising Costs of Training Frontier AI Models," arXiv preprint arXiv:2405.21015 (2024), https://arxiv.org/abs/2405.21015.

[54] In an April 2024 interview, Anthropic CEO Dario Amodei said that within the next two years he expected the cost of training leading models to reach $5 to $10 billion: "Transcript: Ezra Klein Interviews Dario Amodei," *The New York Times*, April 12, 2024, https://www.nytimes.com/2024/04/12/podcasts/transcript-ezra-klein-interviews-dario-amodei.html. For more information on scaling laws and their economic repercussions, see: Drew Lohn, "Scaling AI" (Center for Security and Emerging Technology, 2023), https://cset.georgetown.edu/publication/scaling-ai/; Schmidt et al., "Evaluating Natural Monopoly Conditions in the AI Foundation Model Market"; Corrigan, "America Goes All-in on Big AI"; Varoquaux et al., "Hype, Sustainability, and the Price of the Bigger-Is-Better Paradigm in AI."

[55] For more information on "small models," see: Kyle Miller and Drew Lohn, "Techniques to Make Large Language Models Smaller: An Explainer" (Center for Security and Emerging Technology, October 11, 2023), https://cset.georgetown.edu/publication/techniques-to-make-large-language-models-smaller-an-explainer/; "OpenAI o1 System Card," *OpenAI*, September 12, 2024, https://cdn.openai.com/o1-preview-system-card-20240917.pdf; Will Knight, "Pocket-Sized AI Models Could Unlock a New Era of

Computing," *Wired*, May 23, 2024, https://www.wired.com/story/pocket-sized-ai-models-unlock-new-era-of-computing/; Krystal Hu and Anna Tong, "OpenAI and Others Seek New Path to Smarter AI as Current Methods Hit Limitations," *Reuters*, November 15, 2024, https://www.reuters.com/technology/artificial-intelligence/openai-rivals-seek-new-path-smarter-ai-current-methods-hit-limitations-2024-11-11/; Patrick Tucker, "Is 'Big AI' Beating 'Small AI'—and What Does It Mean for the Military?" *DefenseOne*, October 7, 2024, https://www.defenseone.com/technology/2024/10/big-ai-prevailing-over-small-ai-and-what-does-mean-military/400111/; Martin Juan José Bucher and Marco Martini, "Fine-Tuned 'Small' LLMs (Still) Significantly Outperform Zero-Shot Generative AI Models in Text Classification," arXiv preprint arXiv:2406.08660 (2024), https://arxiv.org/abs/2406.08660.

[56] Charlie Snell, Jaehoon Lee, Kelvin Xu, and Aviral Kumar, "Scaling LLM Test-Time Compute Optimally Can Be More Effective Than Scaling Model Parameters," arXiv preprint arXiv:2408.03314 (2024), https://arxiv.org/pdf/2408.03314.

[57] Brian Merchant, "AI Generate Business" (AI Now Institute, 2024), https://ainowinstitute.org/wp-content/uploads/2024/12/AI-Now_Generative-AI-Business-Models.pdf; Mike Isaac and Erin Griffith, "OpenAI Is Growing Fast and Burning Through Piles of Money," *The New York Times*, September 27, 2024, https://www.nytimes.com/2024/09/27/technology/openai-chatgpt-investors-funding.html; in December 2024, OpenAI introduced a substantially more expensive subscription package: "Introducing ChatGPT Pro," *OpenAI*, December 5, 2024, https://openai.com/index/introducing-chatgpt-pro/.

[58] "Gen AI: Too Much Spend, Too Little Benefit?" Goldman Sachs, June 25, 2024, https://www.goldmansachs.com/images/migrated/insights/pages/gs-research/gen-ai--too-much-spend,-too-little-benefit-/TOM_AI%202.0_ForRedaction.pdf; "What Happened to the Artificial-Intelligence Revolution?" *The Economist*, July 2, 2024, https://www.economist.com/finance-and-economics/2024/07/02/what-happened-to-the-artificial-intelligence-revolution.

[59] Kieran Garvey, Prateek Gupta, Drew Propson et al., "How Agentic AI Will Transform Financial Services With Autonomy, Efficiency and Inclusion," World Economic Forum, December 2, 2024, https://www.weforum.org/stories/2024/12/agentic-ai-financial-services-autonomy-efficiency-and-inclusion/; Shrestha Basu Mallick and Kathy Korevec, "The Next Chapter of the Gemini Era for Developers," *Google Blog*, December 11, 2024, https://developers.googleblog.com/en/the-next-chapter-of-the-gemini-era-for-developers/.

[60] Jack Corrigan, Ngor Luong, and Christian Schoeberl, "Acquiring AI Companies" (Center for Security and Emerging Technology, 2024), https://cset.georgetown.edu/publication/acquiring-ai-companies-tracking-u-s-ai-mergers-and-acquisitions/; Matt O'Brien and Sarah Parvini, "US Senators Call Out Big Tech's New Approach to Poaching Talent, Products from Smaller AI Startups," *Associated Press*, July 12, 2024, https://apnews.com/article/ai-artificial-intelligence-acquihires-amazon-adept-wyden-fa3cd0502a757e5a9ccb83d00bf9ad55.

[61] Schoeberl and Corrigan, "Funding the AI Cloud—Amazon, Alphabet, and Microsoft's Cloud Computing Investments, Part 1"; Schoeberl and Corrigan, "Funding the AI Cloud—Amazon, Alphabet, and Microsoft's Cloud Computing Investments, Part 2"; Christian Schoeberl and Jack Corrigan, "Funding the

AI Cloud—Amazon, Alphabet, and Microsoft's Cloud Computing Investments, Part 3," Center for Security and Emerging Technology (blog), November 20, 2024, https://cset.georgetown.edu/publication/funding-the-ai-cloud-amazon-alphabet-and-microsofts-cloud-computing-investments-part-3/; Corrigan et al., "Acquiring AI Companies."

[62] Subcommittee on Antitrust, Commercial and Administrative Law of the Committee on the Judiciary, *Investigation of Competition in Digital Markets: Majority Staff Report and Recommendations*, https://democrats-judiciary.house.gov/uploadedfiles/competition_in_digital_markets.pdf.

[63] Appenzeller et al., "Navigating the High Cost of AI Compute."

[64] Konstantin Pilz and Lennart Heim, "Compute at Scale," July 2023, https://konstantinpilz.com/data-centers/report/.

[65] Vipra and Myers West, "Computational Power and AI."

[66] Erin Griffith and Cade Metz, "The New A.I. Deal: Buy Everything but the Company," *The New York Times,* August 8, 2024, https://www.nytimes.com/2024/08/08/technology/ai-start-ups-google-microsoft-amazon.html.

[67] Vipra and Myers West, "Computational Power and AI."

[68] Office of Technology, *Partnerships Between Cloud Service Providers and AI Developers*.

[69] Griffith and Metz, "The New A.I. Deal: Buy Everything but the Company"; Corrigan et al., "Acquiring AI Companies."

[70] Narechania and Sitaraman, "An Antimonopoly Approach to Governing Artificial Intelligence"; Office of Technology, *Partnerships Between Cloud Service Providers and AI Developers*.

[71] Complaint, Federal Trade Commission v. Amazon.com, 2:23-cv-01495-JHC, https://www.ftc.gov/system/files/ftc_gov/pdf/1910134amazonecommercecomplaintrevisedredactions.pdf; Michael Liedtke, "Epic Games Wins Antitrust Lawsuit Against Google Over Barriers to its Android App Store," *Associated Press,* December 11, 2023, https://apnews.com/article/google-epic-games-antitrust-trial-android-app-store-dd6b26be7447b5ff8cc0d20a4d01b6b4.

[72] Bjorn Lundqvist, "Cloud Services as the Ultimate Gatekeeper" *Journal of Antitrust Enforcement* 7 (2019), https://www.diva-portal.org/smash/get/diva2:1360698/FULLTEXT01.pdfTest.

[73] Kamila Benzina, "Cloud Infrastructure-as-a-Service as an Essential Facility: Market Structure, Competition, and the Need for Industry and Regulatory Solutions," *Berkeley Technology Law Journal 34,* no. 1 (2019), https://lawcat.berkeley.edu/record/1128982?ln=en&v=pdf; Narechania and Sitaraman, "An Antimonopoly Approach to Governing Artificial Intelligence"; Schoeberl and Corrigan, "Funding the AI Cloud—Amazon, Alphabet, and Microsoft's Cloud Computing Investments, Part 2."

[74] Subcommittee on Antitrust, Commercial and Administrative Law of the Committee on the Judiciary,

Investigation of Competition in Digital Markets: Majority Staff Report and Recommendations; "US FTC Plans to investigate Microsoft's Cloud Business," *Reuters*, November 14, 2024, https://www.reuters.com/technology/us-regulators-plan-investigate-microsofts-cloud-business-ft-reports-2024-11-14/; though some cloud providers, like Amazon and Google, have recently started waiving fees associated with cloud migrations, these policy changes have not fully eliminated the costs of switching providers: David Linthicum, "Why Public Cloud Providers Are Cutting Egress Fees," *InfoWorld*, March 15, 2024, https://www.infoworld.com/article/2336430/why-public-cloud-providers-are-cutting-egress-fees.html; Alex Hawkes, "The Truth About Cloud Data Egress Fees: They're Not Going Away Yet," *PCCW Global*, April 5, 2024, https://blog.consoleconnect.com/the-truth-about-cloud-data-egress-fees.

[75] Narechania and Sitaraman, "An Antimonopoly Approach to Governing Artificial Intelligence."

[76] The exact costs of running inference and model pre-training or fine-tuning depend on a variety of factors, including the amount of inference that is being run and the exact model it is being run on. Furthermore, pre-training foundation models entails high upfront costs, while inference costs continue over time. For more information, see: Cory Weinberg, "OpenAI Projections Imply Losses Tripling to $14 Billion in 2026," *The Information*, October 9, 2024, https://www.theinformation.com/articles/openai-projections-imply-losses-tripling-to-14-billion-in-2026.

[77] Kevin Lee, Adi Gangidi, and Mathew Oldham, "Building Meta's GenAI Infrastructure," *Meta,* March 12, 2024, https://engineering.fb.com/2024/03/12/data-center-engineering/building-metas-genai-infrastructure/.

[78] This includes $23 billion in capital costs and $4 billion in operating costs: Michael Cembalast, "Eye on the Market," JP Morgan, September 3, 2024, https://am.jpmorgan.com/content/dam/jpm-am-aem/global/en/insights/eye-on-the-market/a-severe-case-of-covidia-amv.pdf#page=10.

[79] Helen Li, "Elon Musk's xAI Is on a Hiring Spree for 'AI Tutors'," *Business Insider*, October 16, 2024, https://www.businessinsider.com/elon-musk-xai-hiring-spree-ai-tutors-2024-10; Radsch, "Dismantling AI Data Monopolies Before It's Too Late"; Blake Brittain, "Pulitzer-Winning Authors Join OpenAI, Microsoft Copyright Lawsuit," *Reuters*, December 20, 2023, https://www.reuters.com/legal/pulitzer-winning-authors-join-openai-microsoft-copyright-lawsuit-2023-12-20/; Yiwen Lu, "Digital Media Outlets Sue OpenAI for Copyright Infringement," *The New York Times*, February 28, 2024, https://www.nytimes.com/2024/02/28/technology/openai-copyright-suit-media.html.

[80] Kyle Wiggers, "AI Training Data Has a Price Tag That Only Big Tech Can Afford," *TechCrunch*, June 1, 2024, https://techcrunch.com/2024/06/01/ai-training-data-has-a-price-tag-that-only-big-tech-can-afford/; Katie Paul and Anna Tong, "Inside Big Tech's Underground Race to Buy AI Training Data," *Reuters*, April 5, 2024, https://www.reuters.com/technology/inside-big-techs-underground-race-buy-ai-training-data-2024-04-05/; Anna Tong, Echo Wang, and Martin Coulter, "Exclusive: Reddit in AI Content Licensing Deal With Google," *Reuters*, February 21, 2024, https://www.reuters.com/technology/reddit-ai-content-licensing-deal-with-google-sources-say-2024-02-22/.

[81] Melissa Heikkila, "A Data Bottleneck Is Holding AI Science Back, Says New Nobel winner," *MIT Technology Review*, October 15, 2024, https://www.technologyreview.com/2024/10/15/1105533/a-data-bottleneck-is-holding-ai-science-back-says-new-nobel-winner/; Julia Love and David Alba, "Google Deepens AI Push Into Health Care With Fitbit, Screenings," *Bloomberg*, March 19, 2024, https://www.bloomberg.com/news/articles/2024-03-19/google-deepens-ai-push-into-health-care-with-fitbit-screenings; "Amazon One Medical Launches AI Tools to Help Doctors Get Back to Focusing on Patient Care," Amazon, October 22, 2024, https://www.aboutamazon.com/news/innovation-at-amazon/amazon-one-medical-launches-ai-tools-for-better-patient-care.

[82] Joshua S. Gans, "Market Power in Artificial Intelligence," *National Bureau of Economic Research*, Working Paper 32270, March 2024, https://www.nber.org/system/files/working_papers/w32270/w32270.pdf.

[83] Coen Van Eenbergen, "Microsoft Increases Bing API prices by 1000%, Problems for DuckDuckGo?" *Techzine*, February 17, 2023, https://www.techzine.eu/news/devops/102113/microsoft-increases-bing-api-prices-by-1000-problems-for-duckduckgo/; Yusef Medhi, "Reinventing Search With a New AI-Powered Microsoft Bing and Edge, Your Copilot for the Web," Microsoft Blog, February 7, 2023, https://blogs.microsoft.com/blog/2023/02/07/reinventing-search-with-a-new-ai-powered-microsoft-bing-and-edge-your-copilot-for-the-web/; "DuckDuckGo Launches DuckAssist: a New Feature That Generates Quick Natural Language Summaries in Response to Search Queries," DuckDuckGo, March 8, 2023, https://spreadprivacy.com/duckassist-launch/.

[84] Wiggers, "AI Training Data Has a Price Tag That Only Big Tech Can Afford."

[85] DeepSeek AI, "DeepSeek-R1: Incentivizing Reasoning Capabilities in LLMs," GitHub.com, January 2025, https://github.com/deepseek-ai/DeepSeek-R1/blob/main/DeepSeek_R1.pdf; Cristina Criddle and Eleanor Olcott, "OpenAI Says It Has Evidence China's DeepSeek Used Its Model to Train Competitor," *The Financial Times*, January 28, 2025, https://www.ft.com/content/a0dfedd1-5255-4fa9-8ccc-1fe01de87ea6.

[86] Kyle Wiggers, "The Promise and Perils of Synthetic Data," *TechCrunch*, December 24, 2024, https://techcrunch.com/2024/12/24/the-promise-and-perils-of-synthetic-data/.

[87] Anthropic relies on both Amazon and Alphabet for compute services: Office of Technology, *Partnerships Between Cloud Service Providers and AI Developers.*

[88] Michael Crider, "Google Couldn't Sell AI for Workspace, so It Upped the Price and Bundled It," *PC World*, January 15, 2025, https://www.pcworld.com/article/2578255/google-couldnt-sell-ai-for-workspace-so-it-upped-the-price-and-bundled.html.

[89] Switching costs will vary based on the particular AI models that the developer uses, and will likely be lower for open models: Widder et al., "Open (for Business): Big Tech, Concentrated Power, and the Political Economy of Open AI".

[90] Office of Technology, "Generative AI Raises Competition Concerns," U.S. Federal Trade Commission, June 29, 2023, https://www.ftc.gov/policy/advocacy-research/tech-at-ftc/2023/06/generative-ai-raises-competition-concerns; Office of Technology, *Partnerships Between Cloud Service Providers and AI Developers*.

[91] Widder et al., "Open (for Business): Big Tech, Concentrated Power, and the Political Economy of Open AI"; Miller, "Open Foundation Models: Implications of Contemporary Artificial Intelligence."

[92] National Telecommunications and Information Administration, *Dual-Use Foundation Models with Widely Available Model Weights Report* (Washington, DC: U.S. Department of Commerce, 2024), https://www.ntia.gov/issues/artificial-intelligence/open-model-weights-report.

[93] Widder et al., "Open (For Business): Big Tech, Concentrated Power, and the Political Economy of Open AI"; Schoeberl and Corrigan, "Funding the AI Cloud—Amazon, Alphabet, and Microsoft's Cloud Computing Investments, Part 2."

[94] Devansh, "Mark Zuckerberg: 'Open Source Is Good Business'," *Medium.com*, February 2, 2024, https://machine-learning-made-simple.medium.com/mark-zuckerberg-open-source-is-good-business-c62828f24bf1; it is worth noting that many developers have pushed back on the idea that Meta's Llama models should be considered "open source" at all given the restrictions baked into its community licenses. For more information, see: Jordan Maris, "Meta's LLaMa license is still not Open Source," *Open Source Initiative,* February 18, 2025, https://opensource.org/blog/metas-llama-license-is-still-not-open-source; Ben Cottier, Josh You, Natalia Martemianova, and David Owen, "How Far Behind Are Open Models?" (Epoch AI, 2024), https://epoch.ai/blog/open-models-report.

[95] Memorandum Opinion, United States v. Google, Case No. 20-cv-3010; Zander Arnao, "The Google Search Antitrust Case Is a Triumph for Behavioral Economics," Knight-Georgetown Institute, October 11, 2024, https://kgi.georgetown.edu/research-and-commentary/the-google-search-antitrust-case-is-a-triumph-for-behavioral-economics/.

[96] Memorandum Opinion, United States v. Google, Case No. 20-cv-3010; Jack Corrigan, "What Google's Antitrust Defeat Means for AI," *TIME*, August 29, 2024, https://time.com/7015493/google-antitrust-defeat-ai-monopolies/. GOTTOHERE

[97] U.S. Federal Trade Commission, "FTC Sues Amazon for Illegally Maintaining Monopoly Power," September 26, 2023, https://www.ftc.gov/news-events/news/press-releases/2023/09/ftc-sues-amazon-illegally-maintaining-monopoly-power.

[98] Office of Public Affairs, "Justice Department Sues Apple for Monopolizing Smartphone Markets," U.S. Department of Justice, March 21, 2024, https://www.justice.gov/opa/pr/justice-department-sues-apple-monopolizing-smartphone-markets; Court's Findings of Fact, United States v. Microsoft Corporation, Civil Action No. 98-1233, https://www.justice.gov/atr/us-v-microsoft-courts-findings-fact.

[99] "Use Gemini, Your AI Assistant, on Your Pixel Phone," Google.com, accessed January 2025,

https://support.google.com/pixelphone/answer/15283615; Jared Spataro, "Introducing Microsoft 365 Copilot—Your Copilot for Work," *Microsoft Blog*, March 16, 2023, https://blogs.microsoft.com/blog/2023/03/16/introducing-microsoft-365-copilot-your-copilot-for-work/.

[100] "Amazon and Anthropic Deepen Their Shared Commitment to Advancing Generative AI," Amazon.com, March 27, 2024, https://www.aboutamazon.com/news/company-news/amazon-anthropic-ai-investment; Amazon Bedrock: Model Catalog, Accessed April 2025.

[101] Dominika Kaminska, "Productivity Software Market Share Worldwide," SaasGenius.com, September 5, 2023, https://www.saasgenius.com/statistic/productivity-software-market-share-worldwide/.

[102] Spataro, "Introducing Microsoft 365 Copilot—Your Copilot for Work"; Aparna Pappu, "Google Workspace Extends Gemini Benefits to Millions of Customers," Google Blog, September 24, 2024, https://workspace.google.com/blog/product-announcements/google-workspace-extends-gemini-benefits-to-more-customers.

[103] Van der Vlist et al., "Big AI: Cloud Infrastructure Dependence and the Industrialisation of Artificial Intelligence."

[104] Corrigan, "America Goes All-in on Big AI"; Hanna Dohmen, Jacob Feldgoise, and Charles Kupchan, "The Limits of the China Chip Ban," *Foreign Affairs*, July 24, 2024, https://www.foreignaffairs.com/china/limits-china-chip-ban.

[105] Tejas N. Narechania, "Machine Learning as Natural Monopoly," *Iowa Law Review 107*, Issue 4 (2022), https://ilr.law.uiowa.edu/print/volume-107-issue-4/machine-learning-as-natural-monopoly; Schmid et al., "Evaluating Natural Monopoly Conditions in the AI Foundation Model Market."

[106] NVIDIA Launches DGX Cloud, Giving Every Enterprise Instant Access to AI Supercomputer From a Browser," NVIDIA, March 21, 2023, https://nvidianews.nvidia.com/news/nvidia-launches-dgx-cloud-giving-every-enterprise-instant-access-to-ai-supercomputer-from-a-browser; Eli Tan and Mike Isaac, "The Company Testing Wall Street's Appetite for A.I. Computing Power," *The New York Times*, March 18, 2025, https://www.nytimes.com/2025/03/18/technology/coreweave-wall-street-ai-ipo.html; Timothy Green, "Oracle's AI Success Is Great News for IBM," *The Motley Fool*, September 11, 2024, https://www.fool.com/investing/2024/09/11/oracles-ai-success-is-great-news-for-ibm/.

[107] Narechania and Sitaraman, "An Antimonopoly Approach to Governing Artificial Intelligence."

[108] Arash Massoudi, James Fontanella-Khan, Stephen Morris, and Stefania Palma, "US Regulators Plan to Investigate Microsoft's Cloud Business," *The Financial Times*, November 14, 2024, https://www.ft.com/content/62f361eb-ce52-47c1-9857-878cfe298d54.

[109] Narechania and Sitaraman, "An Antimonopoly Approach to Governing Artificial Intelligence."

[110] Joel Dodge, "Creating a Public Cloud through the Defense Production Act" (Vanderbilt Policy Accelerator, 2024), https://cdn.vanderbilt.edu/vu-URL/wp-content/uploads/sites/412/2024/09/27135238/VPA-Paper-DPA-Public-Cloud.pdf.

[111] "National Artificial Intelligence Research Resource Pilot," National Science Foundation, Accessed January 2025, https://new.nsf.gov/focus-areas/artificial-intelligence/nairr.

[112] Matt Bruenig and Zephyr Teachout, "Should the Government Break Up Big Corporations or Buy Them?," *The Nation*, September 27, 2024, https://www.thenation.com/article/economy/nationalize-corporations-anti-monopoly-debate/.

[113] U.S. Federal Trade Commission, "FTC Launches Inquiry into Generative AI Investments and Partnerships," January 25, 2024, https://www.ftc.gov/news-events/news/press-releases/2024/01/ftc-launches-inquiry-generative-ai-investments-partnerships; Schoeberl and Corrigan, "Funding the AI Cloud—Amazon, Alphabet, and Microsoft's Cloud Computing Investments, Part 1"; Schoeberl and Corrigan, "Funding the AI Cloud—Amazon, Alphabet, and Microsoft's Cloud Computing Investments, Part 2"; Christian Schoeberl and Jack Corrigan, "Funding the AI Cloud—Amazon, Alphabet, and Microsoft's Cloud Computing Investments, Part 3"; Corrigan et al., "Acquiring AI Companies."

[114] "U.S. v. Microsoft: Timeline," *Wired*, November 4, 2002, https://www.wired.com/2002/11/u-s-v-microsoft-timeline/; Ian Bogost, "Google Already Won," *The Atlantic*, August 6, 2024, https://www.theatlantic.com/technology/archive/2024/08/google-monopoly-ruling/679379/.

[115] Noman Bashir, Priya Donti, James Cuff et al., "The Climate and Sustainability Implications of Generative AI," *An MIT Exploration of Generative AI*, March 27, 2024, https://mit-genai.pubpub.org/pub/8ulgrckc/release/2; Tucker, "Is 'Big AI' Beating 'Small AI'—and What Does It Mean for the Military?".

[116] Sukhi Gulati-Gilbert and Robert Seamans, "Data Portability and Interoperability: A Primer on Two Policy Tools for Regulation of Digitized Industries," Brookings Institution, May 9, 2023, https://www.brookings.edu/articles/data-portability-and-interoperability-a-primer-on-two-policy-tools-for-regulation-of-digitized-industries-2/; "The Potential of Data Portability," *Data Transfer Initiative*, 2025, https://dtinit.org/assets/DTIvisionpaperGlobal.pdf.

[117] Narechania and Sitaraman, "An Antimonopoly Approach to Governing Artificial Intelligence."

[118] Digital Markets Act, 2022, (EU) 2022/1925, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32022R1925; American Innovation and Choice Online Act, S. 2992, 117th Cong. (2021), https://www.congress.gov/bill/117th-congress/senate-bill/2992/text.

[119] Memorandum Opinion, United States v. Google, Case No. 20-cv-3010; Corrigan, "What Google's Antitrust Defeat Means for AI."

[120] Will Douglas Heaven, "Small Language Models: 10 Breakthrough Technologies 2025," *MIT Technology Review*, January 3, 2025, https://www.technologyreview.com/2025/01/03/1108800/small-language-models-ai-breakthrough-technologies-2025/.

[121] Jamie A. Sandhu, "Making Big Leaps With Small Models: What Are Small Language Models and Super Tiny Language Models?," Schwartz Reisman Institute for Technology and Society, October 1, 2024, https://srinstitute.utoronto.ca/news/what-are-small-super-tiny-language-models.

[122] Emilia David, "Nvidia's AI Agent Play Is Here With New Models, Orchestration Blueprints," *VentureBeat*, January 6, 2025, https://venturebeat.com/ai/nvidias-ai-agent-play-is-here-with-new-models-orchestration-blueprints/.

[123] National Telecommunications and Information Administration, *Dual-Use Foundation Models with Widely Available Model Weights Report.*

[124] Miller, "Open Foundation Models: Implications of Contemporary Artificial Intelligence."

[125] Anton Korinek and Jai Vipra, "Concentrating Intelligence: Scaling and Market Structure in Artificial Intelligence," *Institute for New Economic Thinking*, Working Paper No. 228, October 2, 2024, https://www.ineteconomics.org/uploads/pjapers/WP_228-Korinek-and-Vipra.pdf.