

July 15, 2024

U.S. Department of Justice Antitrust Division
950 Pennsylvania Avenue NW
Washington, DC 20530

INTRODUCTION AND SUMMARY

The Information Technology and Innovation Foundation (ITIF), the world's leading think tank for science and technology, appreciates the opportunity to contribute to the Department of Justice's (DOJ) Antitrust Division to support its work following the May 30, 2024 workshop on promoting competition in artificial intelligence (AI).¹ This submission represents feedback from ITIF's Center for Data Innovation and ITIF's Schumpeter Project on Competition Policy.

There are three key points we want to highlight in this submission to support the Department in promoting competition in AI:

1. The AI market is amorphous, nascent, dynamic, and competitive;
2. Partnerships between AI firms, including large firms and startups, as well as vertical integration in the AI space, have several procompetitive benefits;
3. Policymakers should adopt a more holistic definition of "openness" that encompasses both system accessibility and the permissibility of AI licenses to better assess potential benefits and risks of models and allow end-users to choose between the benefits of open versus closed AI models.

1. The AI market is amorphous, nascent, dynamic, and competitive.

In general, the AI space is characterized by healthy competition that drives innovation and long-term growth—both from new AI entrants and incumbents. New entrants like OpenAI, Anthropic, Cohere, and Mistral AI are competing vigorously with established tech giants such as Google, Microsoft, Meta, Amazon, and Tencent.

The generative AI market is still in its early stages, and as of now, there is no evidence of significant entry barriers. For example, concerns about data being an entry barrier in AI are speculative and unsubstantiated. Firms seeking to create generative AI models can use data from various sources, including publicly available data on the Internet, government and open-source datasets, datasets licensed from rightsholders, data from workers, and data shared by users. They also have the option to generate synthetic data to train their models.² Some firms, such as OpenAI, Anthropic, and Mistral AI, have succeeded in creating leading generative AI models despite not having access to the large corpus of user data held by social media companies such as Meta and X. Additionally, companies with internal data can leverage it to build specialized models tailored to specific tasks or fields, such as financial services or healthcare.

Similarly, compute resources required for training generative AI models have not proven to be an entry barrier. There are numerous players in the cloud server market that provide the necessary infrastructure for training and running AI models. In some cases, a firm building AI models may even use a rival's cloud services. For example, Anthropic used Google Cloud to train its Claude AI models.³

In terms of chips, Nvidia's graphics processing units (GPUs) are popular but face meaningful potential competition from firms such as AMD and Intel. Indeed, Nvidia itself is a Schumpeterian success story as a firm that leapfrogged chip suppliers like Intel and Qualcomm who were leaders during the computing and mobile technology waves. While concentration in high-tech hardware markets is not surprising given high levels of fixed costs and research and development, other firms are also investing in chip design and manufacturing to prove more efficient, faster, or cheaper chips for certain tasks.⁴ For example, Google has invested heavily in Tensor Processing Units (TPUs), which are specialized chips designed to train and run AI models. Beyond GPUs and TPUs, which are mostly used to train and develop AI algorithms, there are field programmable gate arrays (FPGAs), which are mostly used to apply trained AI algorithms to new data inputs. FPGAs are different from other AI chips because their architecture can be modified by programmers after fabrication. There is also group of AI chips called application-specific integrated circuits (ASICs), which can be used for either training or inference tasks. ASICs have hardware that is customized for a specific algorithm and typically provide more efficiency than FPGAs, but because they are so narrow in their application, they grow obsolete more quickly as new AI algorithms are created. In the long term, there are many areas of AI chips for competitors to thrive, especially those like Infineon that make more energy-efficient chips as the use of electricity is proving to be a huge cost for companies training and running AI models.⁵

Overall, the AI industry is highly dynamic and competitive, allowing new companies to rise to prominence, and challenge established leaders. This ongoing process fosters strong competition at various levels of the AI stack, ensuring a healthy and innovative market.

2. Partnerships between AI firms, including large firms and startups, as well as vertical integration in the AI space, have several procompetitive benefits.

Some firms in the AI market are vertically integrated, meaning they provide multiple components along the AI value chain, such as cloud infrastructure, AI models, and end-user applications. A brewing, yet unsubstantiated, concern is that large, vertically integrated firms may engage in anticompetitive practices to hurt rivals. For example, some worry that a large, vertically integrated firm could restrict access to essential cloud resources or duplicate features from smaller competitors, effectively squeezing them out of the market due to their own larger scale and reach. Additionally, these firms might prioritize their own AI products and services within their ecosystem, limiting market access for new entrants. As a result, several competition authorities would like to see "mix-and-match" competition at and between all layers of the AI stack rather than vertical integration.

However, a market that includes both vertically integrated firms and independent providers can be more competitive because these vertically integrated firms have benefits. For instance, a cloud provider that also

offers an AI model can combine products to lower costs and improve efficiency, leading to more production and better services for end-users. Vertical integration can also cut down on extra costs and encourage innovation, as companies are motivated to keep all parts of their system running smoothly to avoid problems.

In the case of a company that offers its own products and services across the entire stack, this interconnectedness creates stronger incentives for them to innovate and maintain high performance across all components to avoid cascading failures. In the case where two firms have partnered, this interconnectedness incentivises firms to partner with high-performing companies and to synchronize their strategies and innovations closely. The pressure to sustain competitiveness throughout the entire value chain drives robust and comprehensive improvements, enhancing overall market competition. Furthermore, the competition between these vertically integrated ecosystems encourages diverse, innovative solutions, as each ecosystem strives to offer superior integrated services. Consequently, the presence of vertical ecosystems alongside independent providers fosters a dynamic, competitive environment where both integrated solutions and specialized components thrive.

Moreover, concerns have been raised about exclusionary conduct in the context of partnerships between large digital firms and AI startups, such as the partnership between Amazon and Anthropic. However, such partnerships do not in and of themselves harm competition in terms of reduced innovation, consumer welfare, or choice. Take the Amazon and Anthropic partnership, which is facing scrutiny from the UK's Competition and Markets Authority (CMA), as an example.⁶ As the Center explains in its comments to the CMA, this partnership does not prevent competitors from using Anthropic's models or Amazon's cloud services. For example, Google recently announced the availability of Anthropic's enterprise large language models (LLMs) in Google Cloud, including Anthropic's Claude 3 Opus, Claude 3 Sonnet, and Claude 3 Haiku.⁷ Developers can choose which cloud platform best fits their needs. For example, Amazon offers features like Amazon CodeWhisperer, an AI coding companion to help developers increase their productivity.⁸ In addition, Anthropic does not have exclusive access to Amazon's cloud services, which Amazon offers widely and seeks to fully monetize. Indeed, Amazon Bedrock, Amazon's fully managed service for access to foundation models, offers customers access to a broad array of AI models, including from Cohere, Meta, Mistral AI, and Stability AI. The partnership does not lessen Amazon's incentive to enter the AI space: The Anthropic investment is complementary to—not at the expense of—Amazon's massive investments in AI innovation. For example, Amazon also offers access to its own models, such as Amazon Titan, through Amazon Bedrock.⁹ Amazon is also reportedly working on a 2-trillion parameter AI model of its own, which would rival any LLM on the market today.¹⁰

Because vertically integrated AI ecosystems can have procompetitive effects that benefit consumers overall, regulators should base decisions about AI and competition on a detailed understanding of markets, including current and future sources of innovation, and focus on increasing social welfare. The DOJ and FTC's guidelines explain that nonprice terms also matter when evaluating a merger or acquisition, including “reduced product quality, reduced product variety, reduced service, or diminished innovation.”¹¹ Since vertical ecosystems in the AI industry often prioritize differentiation over price competition, emphasizing

offering unique features, innovative solutions, and high-quality services to distinguish themselves in the market, regulators should consider this focus on differentiation when evaluating the competitive landscape of AI ecosystems.

3. Policymakers should adopt a more holistic definition of “openness” that encompasses both system accessibility and the permissibility of AI licenses to better assess potential benefits and risks of models and allow end-users to choose between the benefits of open versus closed AI models.

One point from the workshop that participants agreed on was that open-source models foster innovation and competition. A key message from the workshop speakers to policymakers was to not rush to erect safety guardrails around AI that might harm competition, especially if these guardrails would threaten open AI models.¹² Policymakers should consider using a more comprehensive definition of what is an “open” model because some closed-source systems bring value to innovation commons through different means.

The traditional measure used to define an “open” or “closed” AI model is solely based on how accessible the components of an AI system are to the public or specific users. A popular framework for understanding the accessibility of AI models categorizes them into six levels: fully closed, gradual or staged access, hosted access, cloud-based or API access, downloadable access, and fully open.¹³ In a fully closed system, the system is entirely inaccessible outside the developer organization. In a fully open system, all components of the system, such as training data, model weights, and source code, are accessible and downloadable, allowing unrestricted use and modification. BLOOM, the multilingual open-source language model developed by the BigScience research community, and Cohere’s language model Aya, are both considered fully open by this definition.

However, this measure based solely on aspects related to system access overlooks important economic factors. A more comprehensive measure would include the permissibility of AI licenses, which refer to the legal terms under which an AI model can be used, modified, and distributed, in addition to system access. The structure of an AI license can significantly influence innovation and competition by controlling the flow of knowledge and information into the knowledge commons, affecting the development and enhancement of new AI models by different firms (horizontal competition). Additionally, these licenses regulate access to foundation models and determine conditions for firms that use, modify, or distribute these models at different levels of the supply chain (vertical competition).¹⁴

A comprehensive definition of openness that includes an assessment of how these licenses facilitate the flow of information and innovation can change the ranking of AI foundation models’ openness. A 2024 paper titled “Measuring the Openness of AI Foundation Models” found that when economic factors are considered, the rankings of major models differed significantly from previous rankings based solely on system access. For instance, Cohere’s Aya was considered more open than BigScience’s BLOOM because the former was released with the with a fully permissive Apache 2.0 license while the latter was released with a Responsible AI License (RAIL), which effectively imposes behavioral-use terms on the use of the model.¹⁵ Not only was the ranking orders of major models different, the magnitude of the size of the gaps between

models were different. They note, “When it comes to the size of the gap, our analysis shows that the distinction between so called “open” and “closed” foundation models is not as clear-cut as a purely technical analysis would like to portray...This warrants caution when talking about “open source” models, as most of the AI foundation models rank in the middle to low end of the openness spectrum.”¹⁶

Policymakers should adopt a more holistic definition of "openness" that encompasses both technical accessibility and the permissibility of AI licenses. This approach will better capture the true competitive and innovative potential of AI models. However, they should also recognize that closed models, like OpenAI's, can also offer benefits to consumers as well as spur incentives to innovate by limiting free riding. In other words, an AI space defined by healthy product differentiation between closed and open models will allow for consumers with varying preferences to choose which model they prefer, much the same way as they do in the mobile platform space between Google and Apple.

Sincerely,

Hodan Omaar
Senior Policy Manager
Information Technology and Innovation Foundation

ENDNOTES

¹ U.S. Department of Justice, Antitrust Division, “Workshop on Promoting Competition in Artificial Intelligence,” event, May 30, 2024, <https://www.justice.gov/atr/event/workshop-promoting-competition-artificial-intelligence>.

² Adam Zewe, “Synthetic Data Can Offer Real Performance Improvements,” MIT News, November 3, 2022, <https://news.mit.edu/2022/synthetic-dataai-improvements-1103>.

³ “Anthropic Partners with Google Cloud,” Anthropic, February 3, 2023, <https://www.anthropic.com/news/anthropic-partners-with-google-cloud>.

⁴ Emilia David, “Chip Race: Microsoft, Meta, Google, and Nvidia Battle It Out for AI Chip Supremacy,” The Verge, February 1, 2024, <https://www.theverge.com/2024/2/1/24058186/ai-chips-meta-microsoftgoogle-nvidia>.

⁵ Sharon Goldman, “Can anyone beat Nvidia in AI? Analysts say it’s the wrong question,” *Fortune*, July 2, 2024, <https://fortune.com/2024/07/02/nvidia-competition-ai-chip-gpu-startups-analysts/>.

⁶ Kelvin Chan, “Microsoft and Amazon face scrutiny from UK competition watchdog over recent AI deals,” *Associated Press*, April 24, 2024, <https://apnews.com/article/microsoft-amazon-anthropic-ai-investment-scrutiny-a52f6409fa6ce9b335ab6801b8e84cde>.

⁷ Daniel Castro, “Comments to the Competition and Market Authority Regarding the Amazon-Anthropic Partnership,” Center for Data Innovation, May 10, 2024, <https://www2.datainnovation.org/2024-cma-amazon-anthropic.pdf>.

⁸ “Amazon CodeWhisperer,” AWS, n.d., <https://aws.amazon.com/codewhisperer/>.

⁹ “Amazon Bedrock,” AWS, n.d., <https://aws.amazon.com/bedrock/>.

¹⁰ “Amazon dedicates team to train ambitious AI model codenamed 'Olympus',” Reuters, November 8, 2023, <https://www.reuters.com/technology/amazon-sets-new-team-trains-ambitious-ai-model-codenamed-olympus-sources2023-11-08/>.

¹¹ US. Department of Justice and the Federal Trade Commission, Horizontal Merger Guidelines, August 19, 2010, 2, <https://www.justice.gov/atr/horizontal-merger-guidelines-08192010>

¹² Ibid.

¹³ Irene Solaiman, “The Gradient of Generative AI Release: Methods and Considerations,” Hugging Face (February 2023), <https://arxiv.org/pdf/2302.04844>.

¹⁴ Thibault Schrepel and Jason Potts, “Measuring the Openness of AI Foundation Models: Competition and Policy Implications,” Sciences Po Digital Governance and Sovereignty Chair, working paper, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4827358.

¹⁵ “Cohere for AI Launches Aya, an LLM Covering More Than 100 Languages,” Cohere, February 13, 2024, <https://cohere.com/blog/aya> and Carlos Muñoz Ferrandis et al., “The BigScience RAIL License,” Hugging Face, n.d., <https://bigscience.huggingface.co/blog/the-bigscience-rail-license>.

¹⁶ Schrepel and Potts, “Openness of AI Foundations.”