International Center for Law & Economics

Comments of the International Center for Law & Economics

JFTC Request for Information and Comments Concerning Generative AI and Competition

November 21, 2024

Authored by:

Geoffrey A. Manne (President and Founder, International Center for Law & Economics)

Dirk Auer (Director of Competition Policy, International Center for Law & Economics)

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

Lazar Radic (Senior Competition Policy Scholar, International Center for Law & Economics)

Mario A. Zuniga (Senior Competition Policy Scholar, International Center for Law & Economics)

Executive Summary

We thank the Japan Fair Trade Commission (JFTC) for this invitation to comment (ITC) on Generative AI and Competition.[1] The International Center for Law & Economics (ICLE) is a nonprofit, nonpartisan global research and policy center founded with the goal of building the intellectual foundations for sensible, economically grounded policy. ICLE promotes the use of law & economics methodologies to inform public-policy debates and has longstanding expertise in the evaluation of competition law and policy. ICLE's interest is to ensure that competition law remains grounded in clear rules, established precedent, a record of evidence, and sound economic analysis.

The JFTC recently published a discussion paper on "Generative AI and Competition" ("Discussion Paper")[2] that identifies potential competition issues and asks specific questions for each of them. Those issues are, mainly, connected to potential foreclosure of "essential" inputs for Generative AI production: semiconductors (GPU's), data, and talent. According to the Discussion Paper, so-called "big tech" companies may have the incentives and ability to foreclose those inputs. The Discussion Papers also refers to possible collusive behavior using Generative AI and the "cornering" of highly skilled talent via partnerships.

In these comments, we express the view that, in general, policymakers' current concerns about competition in AI industries, including "Generative AI", may be unwarranted. This is particularly true of the notions that data-network effects shield incumbents in AI markets from competition; that Web 2.0's most successful platforms will be able to leverage their competitive positions to dominate generative-AI markets; that these same platforms may use strategic partnerships with AI firms to insulate themselves from competition; and that generative-AI services occupy narrow markets that leave firms with significant market power.

In fact, we are still far from understanding the boundaries of antitrust-relevant markets in AI. There are three main things that need to be at the forefront of competition authorities' minds when they think about market definition in AI products and services. First, understand that the "AI market" is not unitary, but is instead composed of many distinct goods and services. Second, and relatedly, look beyond the AI marketing hype to see how this extremely heterogeneous products landscape intersects with an equally variegated consumer-demand landscape.

In other words: AI products and services may, in many instances, be substitutable for non-AI products, which would mean that, for the purposes of antitrust law, AI and non-AI products contend in the same relevant market. Getting this relevant product-market definition right is important in antitrust because wrong market definitions could lead to wrong inferences about market power. While

---

[1] Press Release, *Requests for Information and Comments Concerning Generative AI and Competition*, Japan Fair Trade Commission (Oct. 2, 2024), https://www.jftc.go.jp/en/pressreleases/yearly-2024/October/1002.html.

[2] Japan Fair Trade Commission, *Generative AI and Competition (Discussion Paper)* (Oct., 2024), https://www.jftc.go.jp/file/241002DiscussionPaperEN.pdf

either an overly broad or overly narrow market definition could lead to both over and underenforcement, we believe the former currently represents the bigger threat.

Third, overenforcement in the field of generative AI could paradoxically engender the very harms that policymakers are seeking to avert. As we explain in greater detail below, preventing so-called "big tech" firms from competing in AI markets (for example, by threatening competition intervention whenever they forge strategic relationships with AI startups, launch their own generative-AI services, or embed such services in their existing platforms) may thwart an important source of competition and continued innovation. In short, competition in AI markets is important,[3] but trying naïvely to hold incumbent (in adjacent markets) tech firms back, out of misguided fears they will come to dominate the AI space, is likely to do more harm than good. It is essential to acknowledge how little we know about these nascent markets and that the most important priority at the moment is simply to ask the right questions that will lead to sound competition policy.

The comments proceed as follows. Section I debunks the notion that incumbent tech platforms can use their allegedly superior datasets to overthrow competitors in markets for generative AI. Section **Error! Reference source not found.** deals with the risks of possible input foreclosure related to computing power, or GPUs. Section III discusses how policymakers should approach Merger Policy in AI, and specifically, strategic partnerships among tech incumbents and AI startups, including the possible "cornering of specialized talent". Section IV outlines some of the challenges to defining relevant product markets in AI, and suggests how enforcers could navigate the perils of market definition in the nascent, fast-moving world of AI.

I. Anticompetitive Leveraging in AI Markets

Antitrust enforcers have recently expressed concern that incumbent tech platforms may leverage their existing market positions and resources (particularly their vast datasets) to stifle competitive pressure from AI startups. As this section explains, however, these fears appear overblown, as well as underpinned by assumptions about data-network effects that are unlikely to play a meaningful role in generative AI. Instead, the competition interventions that policymakers are contemplating would, paradoxically, remove an important competitive threat for today's most successful AI providers, thereby reducing overall competition in generative-AI markets.

Subsection A summarizes recent calls for competition intervention in generative-AI markets. Subsection BB argues that many of these calls are underpinned by fears of data-related incumbency advantages (often referred to as "data-network effects"), including in the context of mergers. Subsection CC explains why these effects are unlikely to play a meaningful role in generative-AI markets. Subsection D offers five key takeaways to help policymakers better weigh the tradeoffs inherent to competition-enforcement interventions in generative-AI markets.

---

[3] Artificial intelligence is, of course, not a market (at least not a relevant antitrust market). Within the realm of what is called "AI," companies offer myriad products and services, and specific relevant markets would need to be defined before assessing harm to competition in specific cases.

A. Calls for Intervention in AI Markets

It was once (and frequently) said that Google's "data monopoly" was unassailable: "If 'big data' is the oil of the information economy, Google has Standard Oil-like monopoly dominance—and uses that control to maintain its dominant position."[4] Similar claims of data dominance have been attached to nearly all large online platforms, including Facebook (Meta), Amazon, and Uber.[5]

While some of these claims continue even today (for example, "big data" is a key component of the U.S. Department of Justice (DOJ) Google Search and adtech antitrust suits),[6] a shiny new data target has emerged in the form of generative artificial intelligence (AI). The launch of ChatGPT in November 2022, as well as the advent of AI image-generation services like Midjourney and Dall-E, have dramatically expanded the public's conception of what is—and what might be—possible to achieve with generative-AI technologies built on massive datasets.

While these services remain both in the early stages of mainstream adoption and in the throes of rapid, unpredictable technological evolution, they nevertheless already appear to be on the radar of competition policymakers around the world. Several antitrust enforcers appear to believe that, by acting now, they can avoid the "mistakes" that purportedly were made during the formative years of Web 2.0.[7] These mistakes, critics assert, include failing to appreciate the centrality of data in online markets, as well as letting mergers go unchecked and allowing early movers to entrench their market positions.[8] As Federal Trade Commission (FTC) Chair Lina Khan has put it: "we are still reeling

---

[4] Nathan Newman, *Taking on Google's Monopoly Means Regulating Its Control of User Data*, HUFFINGTON POST (Sep. 24, 2013), http://www.huffingtonpost.com/nathan-newman/taking-on-googlesmonopol_b_3980799.html.

[5] *See, e.g.*, Lina Khan & K. Sabeel Rahman, *Restoring Competition in the U.S. Economy*, in UNTAMED: HOW TO CHECK CORPORATE, FINANCIAL, AND MONOPOLY POWER (Nell Abernathy, Mike Konczal, & Kathryn Milani, eds., 2016), at 23. ("From Amazon to Google to Uber, there is a new form of economic power on display, distinct from conventional monopolies and oligopolies..., leverag[ing] data, algorithms, and internet-based technologies... in ways that could operate invisibly and anticompetitively."); Mark Weinstein, *I Changed My Mind–Facebook Is a Monopoly*, WALL ST. J. (Oct. 1, 2021), https://www.wsj.com/articles/facebook-is-monopoly-metaverse-users-advertising-platforms-competition-mewe-big-tech-11633104247 ("[T]he glue that holds it all together is Facebook's monopoly over data.... Facebook's data troves give it unrivaled knowledge about people, governments—and its competitors.").

[6] *See, generally,* Abigail Slater, *Why "Big Data" Is a Big Deal*, THE REG. REV. (Nov. 6, 2023), https://www.theregreview.org/2023/11/06/slater-why-big-data-is-a-big-deal; Amended Complaint at ¶36, *United States v. Google*, 1:20-cv-03010- (D.D.C. 2020); Complaint at ¶37, *United States v. Google*, 1:23-cv-00108 (E.D. Va. 2023), https://www.justice.gov/opa/pr/justice-department-sues-google-monopolizing-digital-advertising-technologies ("Google intentionally exploited its massive trove of user data to further entrench its monopoly across the digital advertising industry.").

[7] *See, e.g.*, Press Release, *Commission Launches Calls for Contributions on Competition in Virtual Worlds and Generative AI*, EUROPEAN COMMISSION (Jan. 9, 2024), https://ec.europa.eu/commission/presscorner/detail/en/IP_24_85; Krysten Crawford, *FTC's Lina Khan Warns Big Tech over AI*, SIEPR (Nov. 3, 2020), https://siepr.stanford.edu/news/ftcs-lina-khan-warns-big-tech-over-ai ("Federal Trade Commission Chair Lina Khan delivered a sharp warning to the technology industry in a speech at Stanford on Thursday: *Antitrust enforcers are watching what you do in the race to profit from artificial intelligence*.") (emphasis added).

[8] *See, e.g.*, John M. Newman, *Antitrust in Digital Markets*, 72 VAND. L. REV. 1497, 1501 (2019) ("[T]he status quo has frequently failed in this vital area, and it continues to do so with alarming regularity. The laissez-faire approach advocated for by scholars and adopted by courts and enforcers has allowed potentially massive harms to go unchecked."); Bertin Martins,

from the concentration that resulted from Web 2.0, and we don't want to repeat the mis-steps of the past with AI."[9]

This response from the competition-policy world is deeply troubling. Rather than engage in critical self-assessment and adopt an appropriately restrained stance, the enforcement community appears to be champing at the bit. Rather than assessing their prior assumptions based on the current technological moment, enforcers' top priority appears to be figuring out how to rapidly and almost reflexively deploy existing competition tools to address the presumed competitive failures presented by generative AI.[10]

It is increasingly common for competition enforcers to argue that so-called "data-network effects" serve not only to entrench incumbents in those markets where the data is collected, but also to confer similar, self-reinforcing benefits in adjacent markets. Several enforcers have, for example, prevented large online platforms from acquiring smaller firms in adjacent markets, citing the risk that they could use their vast access to data to extend their dominance into these new markets.[11]

They have also launched consultations to ascertain the role that data plays in AI competition. For instance, in a recent consultation, the European Commission asked: "What is the role of data and what are its relevant characteristics for the provision of generative AI systems and/or components, including AI models?"[12] Unsurprisingly, the FTC has likewise been hypervigilant about the risks ostensibly posed by incumbents' access to data. In comments submitted to the U.S. Copyright Office, for example, the FTC argued that:

> The rapid development and deployment of AI also poses potential risks to competition. The rising importance of AI to the economy may further lock in the market dominance of large incumbent technology firms. These powerful, vertically integrated incumbents control many of the inputs necessary for the effective development and deployment of AI tools, including cloud-based or local computing power *and access to large stores of training data*. These dominant technology companies may have the incentive to use their

---

Are New EU Data Market Regulations Coherent and Efficient?, Bruegel Working Paper 21/23 (2023), https://www.bruegel.org/working-paper/are-new-eu-data-market-regulations-coherent-and-efficient ("Technical restrictions on access to and re-use of data may result in failures in data markets and data-driven services markets."); Valéria Faure-Muntian, *Competitive Dysfunction: Why Competition Law Is Failing in a Digital World*, THE FORUM NETWORK (Feb. 24, 2021), https://www.oecd-forum.org/posts/competitive-dysfunction-why-competition-law-is-failing-in-a-digital-world.

[9] *See* Rana Foroohar, *The Great US-Europe Antitrust Divide*, FINANCIAL TIMES (Feb. 5, 2024), https://www.ft.com/content/065a2f93-dc1e-410c-ba9d-73c930cedc14.

[10] *See, e.g.*, Press Release, European Commission, *supra* note 7.

[11] *See infra*, Section I.B. Commentators have also made similar claims; *see, e.g.*, Ganesh Sitaram & Tejas N. Narechania, *It's Time for the Government to Regulate AI. Here's How*, POLITICO (Jan. 15, 2024) ("All that cloud computing power is used to train foundation models by having them "learn" from incomprehensibly huge quantities of data. Unsurprisingly, the entities that own these massive computing resources are also the companies that dominate model development. Google has Bard, Meta has LLaMa. Amazon recently invested $4 billion into one of OpenAI's leading competitors, Anthropic. And Microsoft has a 49 percent ownership stake in OpenAI — giving it extraordinary influence, as the recent board struggles over Sam Altman's role as CEO showed.").

[12] Press Release, European Commission, *supra* note 7.

> control over these inputs to unlawfully entrench their market positions in AI and related markets, including digital content markets.[13]

Recently, in the conference that prompts these comments, Jonathan Kanter, assistant U.S. attorney general for antitrust, claimed that:

> We also see structures and trends in AI that should give us pause AI relies on massive amounts of data and computing power, which can give already dominant firms a substantial advantage. Powerful networks and feedback effects may enable dominant firms to control these new markets, and existing power in the digital economy may create a powerful incentive to control emerging innovations that will not only impact our economy, but the health and well-being of our society and free expression itself.[14]

On an even more hyperbolic note, Andreas Mundt, the head of Germany's Federal Cartel Office, called AI a "first-class fire accelerator" for anticompetitive behavior and argued it "will make all the problems only worse."[15] He further argued that "there's a great danger that we'll will get an even deeper concentration of digital markets and power increase at various levels, from chips to the front end."[16] In short, Mundt is one of many policymakers who believes that AI markets will enable incumbent tech firms to further entrench their market positions.

These concerns prompted a joint statement[17] from the FTC, the DOJ, the European Commission, and the United Kingdom's Competition and Market Authority (CMA) in which, while acknowledging that "(a)t their best, these technologies could materially benefit our citizens, boost innovation and drive economic growth," the agencies insist upon "being vigilant and safeguarding against tactics that could undermine fair competition." The statement highlighted three primary risks to competition, summarized below:

1. Concentrated control of key inputs that potentially could put a small number of companies in position to exploit existing or emerging bottlenecks across the AI stack.

2. Entrenching or extending market power in AI-related markets: large incumbent digital firms that already enjoy strong accumulated advantages could use these to protect themselves

---

[13] Comment of U.S. Federal Trade Commission to the U.S. Copyright Office, Artificial Intelligence and Copyright, Docket No. 2023-6 (Oct. 30, 2023), at 4, https://www.ftc.gov/legal-library/browse/advocacy-filings/comment-federal-trade-commission-artificial-intelligence-copyright (emphasis added).

[14] Jonathan Kanter, Remarks at the Promoting Competition in AI Conference (May 30, 2024), https://youtu.be/yh-1AGf3aU?t=424.

[15] Karin Matussek, *AI Will Fuel Antitrust Fires, Big Tech's German Nemesis Warns,* BLOOMBERG (Jun. 26, 2024), https://www.bloomberg.com/news/articles/2024-06-26/ai-will-fuel-antitrust-fires-big-tech-s-german-nemesis-warns?srnd=technology-vp.

[16] *Id.*

[17] Department of Justice, Federal Trade Commission, European Commission & Competition and Markets Authority, *Joint Statement on Competition in Generative AI Foundation Models and AI Products* (Jul. 2024), https://www.ftc.gov/system/files/ftc_gov/pdf/ai-joint-statement.pdf

against AI-driven disruption. This, in turn, may allow such firms to extend or entrench their positions to the detriment of future competition.

3. Arrangements involving key players could amplify risks. Major firms could use partnerships, financial investments, and other connections among firms related to the development of generative to undermine or co-opt competitive threats and steer market outcomes in their favor at the expense of the public.[18]

Certainly, it makes sense that the largest online platforms—including Alphabet, Meta, Apple, and Amazon—should have a meaningful advantage in the burgeoning markets for generative-AI services. After all, it is widely recognized that data is an essential input for generative AI.[19] This competitive advantage should be all the more significant, given that these firms have been at the forefront of AI technology for more than a decade. Over this period, Google's DeepMind and AlphaGo and Meta's NLLB-200 have routinely made headlines.[20] Apple and Amazon also have vast experience with AI assistants, and all of these firms deploy AI technologies throughout their platforms.[21]

Contrary to what one might expect, however, the tech giants have, to date, been largely unable to leverage their vast troves of data to outcompete startups like OpenAI and Midjourney. At the time of writing, OpenAI's ChatGPT appears to be, by far, the most successful chatbot,[22] despite the large tech platforms' apparent access to far more (and more up-to-date) data.

Moreover, it is important not to neglect the role that open-source models currently play in fostering innovation and competition. As former DOJ Chief Antitrust Economist Susan Athey pointed out in a recent interview, "[the AI industry] *may be very concentrated, but if you have two or three high quality – and we have to find out what that means, but high enough quality – open models, then that could be enough*

---

[18] *Id.*

[19] *See, e.g.*, Joe Caserta, Holger Harreis, Kayvaun Rowshankish, Nikhil Srinidhi, & Asin Tavakoli, *The Data Dividend: Fueling Generative AI*, MCKINSEY DIGITAL (Sep. 15, 2023), https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/the-data-dividend-fueling-generative-ai ("Your data and its underlying foundations are the determining factors to what's possible with generative AI.").

[20] *See, e.g.*, Tim Keary, *Google DeepMind's Achievements and Breakthroughs in AI Research*, TECHOPEDIA (Aug. 11, 2023), https://www.techopedia.com/google-deepminds-achievements-and-breakthroughs-in-ai-research; *see, e.g.*, Will Douglas Heaven, *Google DeepMind Used a Large Language Model to Solve an Unsolved Math Problem*, MIT TECHNOLOGY REVIEW (Dec. 14, 2023), https://www.technologyreview.com/2023/12/14/1085318/google-deepmind-large-language-model-solve-unsolvable-math-problem-cap-set; *see also, A Decade of Advancing the State-of-the-Art in AI Through Open Research*, META (Nov. 30, 2023), https://about.fb.com/news/2023/11/decade-of-advancing-ai-through-open-research; *see also, 200 Languages Within a Single AI Model: A Breakthrough in High-Quality Machine Translation*, META, https://ai.meta.com/blog/nllb-200-high-quality-machine-translation (*last visited* Jan. 18, 2023).

[21] *See, e.g.*, Jennifer Allen, *10 Years of Siri: The History of Apple's Voice Assistant*, TECH RADAR (Oct. 4, 2021), https://www.techradar.com/news/siri-10-year-anniversary; *see also* Evan Selleck, *How Apple Is Already Using Machine Learning and AI in iOS*, APPLE INSIDER (Nov. 20, 2023), https://appleinsider.com/articles/23/09/02/how-apple-is-already-using-machine-learning-and-ai-in-ios; *see also*, Kathleen Walch, *The Twenty Year History Of AI At Amazon*, FORBES (Jul. 19, 2019), https://www.forbes.com/sites/cognitiveworld/2019/07/19/the-twenty-year-history-of-ai-at-amazon.

[22] *See infra* Section I.C.

to constrain the for-profit LLMs."[23] Open-source models are important because they allow innovative startups to build upon models already trained on large datasets—therefore entering the market without incurring that initial cost. Apparently, there is no lack of open-source models, since companies like xAI, Meta, and Google offer their AI models for free.[24]

There are important lessons to glean from these developments, if only enforcers would stop to reflect. The meteoric rise of consumer-facing AI services should offer competition enforcers and policymakers an opportunity for *introspection*. As we explain, the rapid emergence of generative-AI technology may *undercut* many core assumptions of today's competition-policy debates, which have focused largely on the rueful after-effects of the purported failure of 20th-century antitrust to address the allegedly manifest harms of 21st-century technology. These include the notions that data advantages constitute barriers to entry and can be leveraged to project dominance into adjacent markets; that scale itself is a market failure to be addressed by enforcers; and that the use of consumer data is inherently harmful to those consumers.

B. Data-Network Effects Theory and Enforcement

Proponents of more extensive intervention by competition enforcers into digital markets often cite data-network effects as a source of competitive advantage and barrier to entry (though terms like "economies of scale and scope" may offer more precision).[25] The crux of the argument is that "the collection and use of data creates a feedback loop of more data, which ultimately insulates incumbent platforms from entrants who, but for their data disadvantage, might offer a better product."[26] This self-reinforcing cycle purportedly leads to market domination by a single firm. Thus, it is argued, *e.g.*, that Google's "ever-expanding control of user personal data, and that data's critical value to online advertisers, creates an *insurmountable barrier to entry for new competition.*"[27]

---

[23] Josh Sisco, *POLITICO PRO Q&A: Exit interview with DOJ Chief Antitrust Economist Susan Athey*, POLITICO PRO (Jul. 2, 2024), https://subscriber.politicopro.com/article/2024/07/politico-pro-q-a-exit-interview-with-doj-chief-antitrust-economist-susan-athey-00166281.

[24] Belle Lin, *Open-Source Companies Are Sharing Their AI Free. Can They Crack OpenAI's Dominance?*, WALL ST. J. (Mar. 21, 2024), https://www.wsj.com/articles/open-source-companies-are-sharing-their-ai-free-can-they-crack-openais-dominance-26149e9c.

[25] *See, e.g.*, Cédric Argenton & Jens Prüfer, *Search Engine Competition with Network Externalities*, 8 J. COMP. L. & ECON. 73, 74 (2012).

[26] John M. Yun, *The Role of Big Data in Antitrust,* in THE GLOBAL ANTITRUST INSTITUTE REPORT ON THE DIGITAL ECONOMY (Joshua D. Wright & Douglas H. Ginsburg, eds., Nov. 11, 2020) at 233, https://gaidigitalreport.com/2020/08/25/big-data-and-barriers-to-entry/#_ftnref50; *see also, e.g.,* Robert Wayne Gregory, Ola Henfridsson, Evgeny Kaganer, & Harris Kyriakou, *The Role of Artificial Intelligence and Data Network Effects for Creating User Value*, 46 ACAD. OF MGMT. REV. 534 (2020), final pre-print version at 4, http://wrap.warwick.ac.uk/134220) ("A platform exhibits data network effects if, the more that the platform learns from the data it collects on users, the more valuable the platform becomes to each user."); *see also,* Karl Schmedders, José Parra-Moyano, & Michael Wade, *Why Data Aggregation Laws Could be the Answer to Big Tech Dominance,* SILICON REPUBLIC (Feb. 6, 2024), https://www.siliconrepublic.com/enterprise/data-ai-aggregation-laws-regulation-big-tech-dominance-competition-antitrust-imd.

[27] Nathan Newman, *Search, Antitrust, and the Economics of the Control of User Data*, 31 YALE J. REG. 401, 409 (2014) (emphasis added); *see also id.* at 420 & 423 ("While there are a number of network effects that come into play with Google, ["its

But it is important to note the conceptual problems these claims face. Because data can be used to improve products' quality and/or to subsidize their use, if possessing data constitutes an entry barrier, then *any* product improvement or price reduction made by an incumbent could be problematic. This is tantamount to an argument that *competition itself* is a cognizable barrier to entry. Of course, it would be a curious approach to antitrust if competition were treated as a problem, as it would imply that firms should under-compete—*i.e.*, should forego consumer-welfare enhancements—in order to inculcate a greater number of firms in a given market, simply for its own sake.[28]

Meanwhile, actual economic studies of data-network effects have been few and far between, with scant empirical evidence to support the theory.[29] Andrei Hagiu and Julian Wright's theoretical paper offers perhaps the most comprehensive treatment of the topic to date.[30] The authors ultimately conclude that data-network effects can be of differing magnitudes and have varying effects on firms' incumbency advantage.[31] They cite Grammarly (an AI writing-assistance tool) as a *potential* example: "As users make corrections to the suggestions offered by Grammarly, its language experts and artificial intelligence can use this feedback to continue to improve its future recommendations for all users."[32]

This is echoed by economists who contend that "[t]he algorithmic analysis of user data and information might increase incumbency advantages, creating lock-in effects among users and making them more reluctant to join an entrant platform."[33] Crucially, some scholars take this logic a step further, arguing that platforms may use data from their "origin markets" in order to enter and dominate adjacent ones:

> First, as we already mentioned, data collected in the origin market can be used, once the enveloper has entered the target market, to provide products more efficiently in the target market. Second, data collected in the origin market can be used to reduce the asymmetric information to which an entrant is typically subject when deciding to invest (for example, in R&D) to enter a new market. For instance, a search engine could be able to

---

intimate knowledge of its users contained in its vast databases of user personal data"] is likely the most important one in terms of entrenching the company's monopoly in search advertising.... Google's overwhelming control of user data... might make its dominance nearly unchallengeable.").

[28] *See also* Yun, *supra* note 26 at 229 ("[I]nvestments in big data can create competitive distance between a firm and its rivals, including potential entrants, but this distance is the result of a competitive desire to improve one's product.").

[29] For a review of the literature on *increasing returns to scale in data* (this topic is broader than data-network effects) *see* Geoffrey Manne & Dirk Auer, *Antitrust Dystopia and Antitrust Nostalgia: Alarmist Theories of Harm in Digital Markets and Their Origins*, 28 GEO MASON L. REV. 1281, 1344 (2021).

[30] Andrei Hagiu & Julian Wright, *Data-Enabled Learning, Network Effects, and Competitive Advantage*, 54 RAND J. ECON. 638 (2023).

[31] *Id.* at 639. The authors conclude that "Data-enabled learning would seem to give incumbent firms a competitive advantage. But how strong is this advantage and how does it differ from that obtained from more traditional mechanisms... ."

[32] *Id.*

[33] Bruno Jullien & Wilfried Sand-Zantman, *The Economics of Platforms: A Theory Guide for Competition Policy*, 54 INFO. ECON. & POL'Y 10080, 101031 (2021).

> predict new trends from consumer searches and therefore face less uncertainty in product design.[34]

This possibility is also implicit in Hagiu and Wright's paper.[35] Indeed, the authors' theoretical model rests on an important distinction between "within-user" data advantages (that is, having access to more data about a given user) and "across-user" data advantages (information gleaned from having access to a wider user base). In both cases, there is an implicit assumption that platforms may use data from one service to gain an advantage in another market (because what matters is information about aggregate or individual user preferences, regardless of its origin).

Our review of the economic evidence suggests that several scholars have, with varying degrees of certainty, raised the possibility that incumbents may leverage data advantages to stifle competitors in their primary market or in adjacent ones (be it via merger or organic growth). As we explain below, however, there is ultimately little evidence to support such claims. Policymakers have nonetheless been keenly receptive to these limited theoretical findings, basing multiple decisions on these theories, often with little consideration given to the caveats that accompany them.[36]

Indeed, it is remarkable that, in its section on "[t]he data advantage for incumbents," the "Furman Report" created for the UK government cited only two empirical economic studies, and they offer *directly contradictory conclusions* with respect to the question of the strength of data advantages.[37] The report nevertheless concluded that data "may confer a form of unmatchable advantage on the incumbent business, making successful rivalry less likely,"[38] and it adopted without reservation what it deemed "convincing" evidence from non-economists that have no apparent empirical basis.[39]

In the *Google/Fitbit* merger proceedings, the European Commission found that the combination of data from Google services with that of Fitbit devices would reduce competition in advertising markets:

---

[34] Daniele Condorelli & Jorge Padilla, *Harnessing Platform Envelopment in the Digital World*, 16 J. COMP. L. & POL'Y 143, 167 (2020).

[35] *See* Hagiu & Wright, *supra* note 30.

[36] For a summary of these limitations, *see generally* Catherine Tucker, *Network Effects and Market Power: What Have We Learned in the Last Decade?*, ANTITRUST (2018) at 72, *available at* https://sites.bu.edu/tpri/files/2018/07/tucker-network-effects-antitrust2018.pdf; *see also* Manne & Auer, *supra* note 29, at 1330.

[37] *See* Jason Furman, Diane Coyle, Amelia Fletcher, Derek McAuley, & Philip Marsden (Dig. Competition Expert Panel), *Unlocking Digital Competition* (2019) at 32-35 ("Furman Report"), *available at* https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/785547/unlocking_digital_competition_furman_review_web.pdf.

[38] *Id.* at 34.

[39] *Id.* at 35. To its credit, it should be noted, the Furman Report does counsel caution before mandating access to data as a remedy to promote competition. *See id.* at 75. That said, the Furman Report maintains that such a remedy should remain on the table because "the evidence suggests that large data holdings are at the heart of the potential for some platform markets to be dominated by single players and for that dominance to be entrenched in a way that lessens the potential for competition for the market." *Id.* The evidence, however, does not show this.

> Giving [sic] the large amount of data already used for advertising purposes that Google holds, the increase in Google's data collection capabilities, which goes beyond the mere number of active users for which Fitbit has been collecting data so far, the Transaction is likely to have a negative impact on the development of an unfettered competition in the markets for online advertising.[40]

As a result, the Commission cleared the merger only on the condition that Google refrain from using data from Fitbit devices for its advertising platform.[41] The Commission also appears likely to focus on similar issues in its ongoing investigation of Microsoft's investment in OpenAI.[42]

Along similar lines, in its complaint to enjoin Meta's purchase of Within Unlimited—makers of the virtual-reality (VR) fitness app Supernatural—the FTC relied on, among other things, the fact that Meta could leverage its data about VR-user behavior to inform its decisions and potentially outcompete rival VR-fitness apps: "Meta's control over the Quest platform also gives it unique access to VR user data, which it uses to inform strategic decisions."[43]

The DOJ's twin cases against Google also implicate data leveraging and data barriers to entry. The agency's adtech complaint charges that "Google intentionally exploited its massive trove of user data to further entrench its monopoly across the digital advertising industry."[44] Similarly, in its Google Search complaint, the agency argued that:

> Google's anticompetitive practices are especially pernicious because they deny rivals scale to compete effectively. General search services, search advertising, and general search text advertising require complex algorithms that are constantly learning which organic results and ads best respond to user queries; the volume, variety, and velocity of data accelerates the automated learning of search and search advertising algorithms.[45]

Finally, updated merger guidelines published in recent years by several competition enforcers cite the acquisition of data as a potential source of competition concerns. For instance, the FTC and DOJ's 2023 guidelines state that "acquiring data that helps facilitate matching, sorting, or prediction services may enable the platform to weaken rival platforms by denying them that data."[46] Likewise,

---

[40] Case COMP/M.9660 – Google/Fitbit, Commission Decision (Dec. 17, 2020) (Summary at O.J. (C 194) 7), *available at* https://ec.europa.eu/competition/mergers/cases1/202120/m9660_3314_3.pdf, at 455,

[41] *Id.* at 896.

[42] *See* Natasha Lomas, *EU Checking if Microsoft's OpenAI Investment Falls Under Merger Rules*, TECHCRUNCH (Jan. 9, 2024), https://techcrunch.com/2024/01/09/openai-microsoft-eu-merger-rules.

[43] Amended Complaint at 11, *Meta/Zuckerberg/Within*, FED. TRADE COMM'N. (2022) (No. 605837), *available at* https://www.ftc.gov/system/files/ftc_gov/pdf/D09411%20-%20AMENDED%20COMPLAINT%20FILED%20BY%20COUNSEL%20SUPPORTING%20THE%20COMPLAINT%20-%20PUBLIC%20%281%29_0.pdf.

[44] Amended Complaint (D.D.C), *supra* note 6 at ¶37.

[45] Amended Complaint (E.D. Va), *supra* note 6 at ¶8.

[46] *Merger Guidelines*, US DEP'T OF JUSTICE & FED. TRADE COMM'N (2023) at 25, *available at* https://www.ftc.gov/system/files/ftc_gov/pdf/2023_merger_guidelines_final_12.18.2023.pdf.

the UK Competition and Markets Authority warned against incumbents acquiring firms in order to obtain their data and foreclose other rivals:

> Incentive to foreclose rivals…
>
> 7.19(e) Particularly in complex and dynamic markets, firms may not focus on short term margins but may pursue other objectives to maximise their long-run profitability, which the CMA may consider. This may include… obtaining access to customer data….[47]

In short, competition authorities around the globe have taken an increasingly aggressive stance on data-network effects. Among the ways this has manifested is in enforcement decisions based on fears that data collected by one platform might confer decisive competitive advantages in adjacent markets. Unfortunately, these concerns rest on little to no empirical evidence, either in the economic literature or the underlying case records.

C. Data-Incumbency Advantages in Generative-AI

Given the assertions detailed in the previous section, it would be reasonable to assume that firms such as Google, Meta, and Amazon should be in pole position to meet the burgeoning demand for generative AI. After all, these firms have not only been at the forefront of the field for the better part of a decade, but they also have access to vast troves of data, the likes of which their rivals could only dream when they launched their own services. Thus, the authors of the Furman Report caution that "to the degree that the next technological revolution centres around artificial intelligence and machine learning, then the companies most able to take advantage of it may well be the existing large companies because of the importance of data for the successful use of these tools."[48]

To date, however, this is not how things have unfolded (although it bears noting that these technologies remain in flux and the competitive landscape is susceptible to change). The first significantly successful generative-AI service was arguably not from either Meta—which had been working on chatbots for years and had access to, arguably, the world's largest database of actual chats—or Google. Instead, the breakthrough came from a previously unknown firm called OpenAI.

OpenAI's ChatGPT service currently accounts for an estimated 60% of visits to online AI tools (though reliable numbers are somewhat elusive).[49] It broke the record for the fastest online service to reach 100 million users (in only a couple of months), more than four times faster than TikTok,

---

[47] *Merger Assessment Guidelines*, COMPETITION AND MKTS. AUTH (2021) at ¶7.19(e), *available at* https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1051823/MAGs_for_publication_2021_~_.pdf.

[48] Furman Report, *supra* note 37, at ¶4.

[49] *See, e.g.*, Chris Westfall, *New Research Shows ChatGPT Reigns Supreme in AI Tool Sector*, FORBES (Nov. 16, 2023), https://www.forbes.com/sites/chriswestfall/2023/11/16/new-research-shows-chatgpt-reigns-supreme-in-ai-tool-sector/?sh=7de5de250e9c; Sujan Sarkar, *AI Industry Analysis: 50 Most Visited AI Tools and Their 24B+ Traffic Behavior*, WRITERBUDDY (*last visited*, Jul. 15, 2024), https://writerbuddy.ai/blog/ai-industry-analysis.

the previous record holder.[50] Based on Google Trends data, ChatGPT is nine times more popular worldwide than Google's own Bard service, and 14 times more popular in the United States.[51] In April 2023, ChatGPT reportedly registered 206.7 million unique visitors, compared to 19.5 million for Google's Bard.[52] In short, at the time we are writing, ChatGPT appears to be the most popular chatbot. The entry of large players such as Google Bard or Meta AI appear to have had little effect thus far on its leading position.[53]

The picture is similar in the field of AI-image generation. As of August 2023, Midjourney, Dall-E, and Stable Diffusion appear to be the three market leaders in terms of user visits.[54] This is despite competition from the likes of Google and Meta, who arguably have access to unparalleled image and video databases by virtue of their primary platform activities.[55]

This raises several crucial questions: how have these AI upstarts managed to be so successful, and is their success just a flash in the pan before Web 2.0 giants catch up and overthrow them? While we cannot answer either of these questions dispositively, we offer what we believe to be some relevant observations concerning the role and value of data in digital markets.

A first important observation is that empirical studies suggest that data exhibits diminishing marginal returns. In other words, past a certain point, acquiring more data does not confer a meaningful edge to the acquiring firm. As Catherine Tucker put it, following a review of the literature: "Empirically there is little evidence of economies of scale and scope in digital data in the instances where one would expect to find them."[56]

---

[50] *See* Krystal Hu, *ChatGPT Sets Record for Fastest-Growing User Base*, REUTERS (Feb. 2, 2023), https://www.reuters.com/technology/chatgpt-sets-record-fastest-growing-user-base-analyst-note-2023-02-01; *Google: The AI Race Is On*, APP ECONOMY INSIGHTS (Feb. 7, 2023), https://www.appeconomyinsights.com/p/google-the-ai-race-is-on.

[51] *See* Google Trends, https://trends.google.com/trends/explore?date=today%205-y&q=%2Fg%2F11khcfz0y2,%2Fg%2F11ts49p01g&hl=en (*last visited* Jan. 12, 2024) and https://trends.google.com/trends/explore?date=today%205-y&geo=US&q=%2Fg%2F11khcfz0y2,%2Fg%2F11ts49p01g&hl=en (*last visited* Jan. 12, 2024).

[52] *See* David F. Carr, *As ChatGPT Growth Flattened in May, Google Bard Rose 187%*, SIMILARWEB BLOG (Jun. 5, 2023), https://www.similarweb.com/blog/insights/ai-news/chatgpt-bard.

[53] *See* Press Release, *Introducing New AI Experiences Across Our Family of Apps and Devices*, META (Sep. 27, 2023), https://about.fb.com/news/2023/09/introducing-ai-powered-assistants-characters-and-creative-tools; Sundar Pichai, *An Important Next Step on Our AI Journey*, GOOGLE KEYWORD BLOG (Feb. 6, 2023), https://blog.google/technology/ai/bard-google-ai-search-updates.

[54] *See* Ion Prodan, *14 Million Users: Midjourney's Statistical Success*, YON (Aug. 19, 2023), https://yon.fun/midjourney-statistics; *see also* Andrew Wilson, *Midjourney Statistics: Users, Polls, & Growth [Oct 2023]*, APPROACHABLEAI (Oct. 13, 2023), https://approachableai.com/midjourney-statistics.

[55] *See* Hema Budaraju, *New Ways to Get Inspired with Generative AI in Search*, GOOGLE KEYWORD BLOG (Oct. 12, 2023), https://blog.google/products/search/google-search-generative-ai-october-update; *Imagine with Meta AI*, META (last visited Jan. 12, 2024), https://imagine.meta.com.

[56] Catherine Tucker, *Digital Data, Platforms and the Usual [Antitrust] Suspects: Network Effects, Switching Costs, Essential Facility*, 54 REV. INDUS. ORG. 683, 686 (2019).

Likewise, following a survey of the empirical literature on this topic, Geoffrey Manne and Dirk Auer conclude that:

> Available evidence suggests that claims of "extreme" returns to scale in the tech sector are greatly overblown. Not only are the largest expenditures of digital platforms unlikely to become proportionally less important as output increases, but empirical research strongly suggests that even data does not give rise to increasing returns to scale, despite routinely being cited as the source of this effect.[57]

In other words, being the firm with the *most* data appears to be far less important than having *enough* data. Moreover, this lower bar may be accessible to far more firms than one might initially think possible. Furthermore, obtaining sufficient data could become easier still—that is, the volume of required data could become even smaller—with technological progress. For instance, synthetic data may provide an adequate substitute to real-world data,[58] or may even outperform real-world data.[59] As Thibault Schrepel and Alex Pentland surmise:

> [A]dvances in computer science and analytics are making the amount of data less relevant every day. In recent months, important technological advances have allowed companies with small data sets to compete with larger ones.[60]

Indeed, past a certain threshold, acquiring more data might not meaningfully improve a service, where other improvements (such as better training methods or data curation) could have a large impact. In fact, there is some evidence that excessive data impedes a service's ability to generate results appropriate for a given query: "[S]uperior model performance can often be achieved with smaller, high-quality datasets than massive, uncurated ones. Data curation ensures that training datasets are devoid of noise, irrelevant instances, and duplications, thus maximizing the efficiency of every training iteration."[61]

Consider, for instance, a user who wants to generate an image of a basketball. Using a model trained on an indiscriminate range and number of public photos in which a basketball appears surrounded by copious other image data, the user may end up with an inordinately noisy result. By contrast, a

---

[57] Manne & Auer, *supra* note 29, at 1345.

[58] *See, e.g.*, Stefanie Koperniak, *Artificial Data Give the Same Results as Real Data—Without Compromising Privacy*, MIT NEWS (Mar. 3, 2017), https://news.mit.edu/2017/artificial-data-give-same-results-as-real-data-0303 ("[Authors] describe a machine learning system that automatically creates synthetic data—with the goal of enabling data science efforts that, due to a lack of access to real data, may have otherwise not left the ground. While the use of authentic data can cause significant privacy concerns, this synthetic data is completely different from that produced by real users—but can still be used to develop and test data science algorithms and models.").

[59] *See, e.g.*, Rachel Gordon, *Synthetic Imagery Sets New Bar in AI Training Efficiency*, MIT NEWS (Nov. 20, 2023), https://news.mit.edu/2023/synthetic-imagery-sets-new-bar-ai-training-efficiency-1120 ("By using synthetic images to train machine learning models, a team of scientists recently surpassed results obtained from traditional 'real-image' training methods.).

[60] Thibault Schrepel & Alex 'Sandy' Pentland, *Competition Between AI Foundation Models: Dynamics and Policy Recommendations*, MIT CONNECTION SCIENCE WORKING PAPER (Jun. 2023), at 8.

[61] Igor Susmelj, *Optimizing Generative AI: The Role of Data Curation*, LIGHTLY (*last visited* Jan. 15, 2024), https://www.lightly.ai/post/optimizing-generative-ai-the-role-of-data-curation.

model trained with a better method on fewer, more carefully selected images could readily yield far superior results.[62] In one important example:

> The model's performance is particularly remarkable, given its small size. "This is not a large language model trained on the whole Internet; this is a relatively small transformer trained for these tasks," says Armando Solar-Lezama, a computer scientist at the Massachusetts Institute of Technology, who was not involved in the new study.... The finding implies that instead of just shoving ever more training data into machine-learning models, a complementary strategy might be to offer AI algorithms the equivalent of a focused linguistics or algebra class.[63]

Platforms' current efforts are thus focused on improving the mathematical and logical reasoning of large language models (LLMs), rather than maximizing training datasets.[64] Two points stand out. The first is that firms like OpenAI rely largely on *publicly available* datasets—such as GSM8K—to train their LLMs.[65] Second, the real challenge to creating innovative AI lies not so much in collecting data, but in creating innovative AI-training processes and architectures:

> [B]uilding a truly general reasoning engine will require a more fundamental architectural innovation. What's needed is a way for language models to learn new abstractions that go beyond their training data and have these evolving abstractions influence the model's choices as it explores the space of possible solutions.
>
> We know this is possible because the human brain does it. But it might be a while before OpenAI, DeepMind, or anyone else figures out how to do it in silicon.[66]

Furthermore, it is worth noting that the data most relevant to startups in a given market may not be those held by large incumbent platforms in other markets. They might instead be data specific to the market in which the startup is active or, even better, to the given problem it is attempting to solve:

> As Andres Lerner has argued, if you wanted to start a travel business, the data from Kayak or Priceline would be far more relevant. Or if you wanted to start a ride-sharing business, data from cab companies would be more useful than the broad, market-cross-cutting profiles Google and Facebook have. Consider companies like Uber, Lyft and

---

[62] *See, e.g.*, Xiaoliang Dai, *et al.*, *Emu: Enhancing Image Generation Models Using Photogenic Needles in a Haystack*, ARXIV (Sep. 27, 2023) at 1, https://ar5iv.labs.arxiv.org/html/2309.15807 ("[S]upervised fine-tuning with a set of surprisingly small but extremely visually appealing images can significantly improve the generation quality."); *see also*, Hu Xu, *et al.*, *Demystifying CLIP Data*, ARXIV (Sep. 28, 2023), https://arxiv.org/abs/2309.16671.

[63] Lauren Leffer, *New Training Method Helps AI Generalize like People Do*, SCI. AM. (Oct. 26, 2023), https://www.scientificamerican.com/article/new-training-method-helps-ai-generalize-like-people-do (discussing Brendan M. Lake & Marco Baroni, *Human-Like Systematic Generalization Through a Meta-Learning Neural Network*, 623 NATURE 115 (2023)).

[64] Timothy B. Lee, *The Real Research Behind the Wild Rumors about OpenAI's Q* Project*, ARS TECHNICA (Dec. 8, 2023), https://arstechnica.com/ai/2023/12/the-real-research-behind-the-wild-rumors-about-openais-q-project.

[65] *Id.*; *see also* GSM8K, PAPERS WITH CODE (*last visited* Jan. 18, 2023), https://paperswithcode.com/dataset/gsm8k; MATH Dataset, GITHUB (*last visited* Jan. 18, 2024), https://github.com/hendrycks/math.

[66] Lee, *supra* note 64.

> Sidecar that had no customer data when they began to challenge established cab companies that *did* possess such data. If data were really so significant, they could never have competed successfully. But Uber, Lyft and Sidecar have been able to effectively compete because they built products that users wanted to use—they came up with an idea for a better mousetrap. The data they have accrued came *after* they innovated, entered the market and mounted their successful challenges—not before.[67]

This point maybe especially relevant in the Japanese context where, as the Discussion Paper emphasizes "Japanese data is crucial for [generative AI products] used in Japan".[68] Here Japanese companies have an advantage over foreign companies that have been modeled and trained with large datasets in English or other languages. Considering that more than half companies in Japan are using generative AI,[69] it is expected that there will be demand for products designed specifically for Japanese markets (with Japanese data), which offers an undeniable growth opportunity for local firms and gives them an advantage, at least in certain market niches, over large, global firms.

The bottom line is that data is not the be-all and end-all that many in competition circles make it out to be. While data may often confer marginal benefits, there is little evidence that these benefits are ultimately decisive.[70] As a result, incumbent platforms' access to vast numbers of users and troves of data in their primary markets might only marginally affect their competitiveness in AI markets.

A related observation is that firms' capabilities and other features of their products arguably play a more important role than the data they own.[71] Examples of this abound in digital markets. Google overthrew Yahoo in search, despite initially having access to far fewer users and far less data. Google and Apple overcame Microsoft in the smartphone operating-system market, despite having comparatively tiny ecosystems (at the time) to leverage. TikTok rose to prominence despite intense competition from incumbents like Instagram, which had much larger userbases. In each of these cases, important product-design decisions (such as the PageRank algorithm, recognizing the specific needs

---

[67] Geoffrey Manne & Ben Sperry, *Debunking the Myth of a Data Barrier to Entry for Online Services*, TRUTH ON THE MARKET (Mar. 26, 2015), https://truthonthemarket.com/2015/03/26/debunking-the-myth-of-a-data-barrier-to-entry-for-online-services (citing Andres V. Lerner, *The Role of 'Big Data' in Online Platform Competition* (Aug. 26, 2014), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2482780.).

[68] Discussion Paper, *supra* note 2, at 6.

[69] Staffing Industry Analysts, *Majority of companies in Japan are using generative AI* (Aug. 28, 2024), https://www.staffingindustry.com/news/global-daily-news/majority-of-companies-in-japan-are-using-generative-ai

[70] *See* Catherine Tucker, *Digital Data as an Essential Facility: Control*, CPI ANTITRUST CHRON. (Feb. 2020), at 11 ("[U]ltimately the value of data is not the raw manifestation of the data itself, but the ability of a firm to use this data as an input to insight.").

[71] Or, as John Yun put it, data is only a small component of digital firms' production function. *See* Yun, *supra* note 26, at 235 ("Second, while no one would seriously dispute that having more data is better than having less, the idea of a data-driven network effect is focused too narrowly on a single factor improving quality. As mentioned in *supra* Section I.A, there are a variety of factors that enter a firm's production function to improve quality.").

of mobile users,[72] and TikTok's clever algorithm) appear to have played far more significant roles than the firms' initial user and data endowments (or lack thereof).

All of this suggests that the early success of OpenAI likely has more to do with its engineering decisions than with what data it did or did not possess. Going forward, OpenAI and its rivals' relative abilities to offer and monetize compelling use cases by offering custom versions of their generative-AI technologies will arguably play a much larger role than (and contribute to) their ownership of data.[73] In other words, the ultimate challenge is arguably to create a valuable platform, of which data ownership is a consequence, not a cause.

It is also important to note that, in those instances where it is valuable, data does not just fall from the sky. Instead, it is through smart business and engineering decisions that firms can generate valuable *information* (which does not necessarily correlate with owning more *data*). For instance, OpenAI's success with ChatGPT is often attributed to its more efficient algorithms and training models, which arguably have enabled the service to improve more rapidly than its rivals.[74] Likewise, the ability of firms like Meta and Google to generate valuable data for advertising arguably depends more on design decisions that elicit the right data from users, rather than the raw number of users in their networks.

Put differently, setting up a business so as to gather and organize the right *information* is more important than simply owning vast troves of *data*.[75] Even in those instances where high-quality data is an essential parameter of competition, it does not follow that having vaster databases or more users on a platform necessarily leads to better information for the platform. Indeed, if data ownership consistently conferred a significant competitive advantage, these new AI firms would not be where they are today.

This does not, of course, mean that data is worthless. Rather, it means that competition authorities should not assume that the mere possession of data is a dispositive competitive advantage, absent compelling empirical evidence to support such a finding. In this light, the current wave of decisions and competition-policy pronouncements that rely on data-related theories of harm are premature.

---

[72] Luxia Le, *The Real Reason Windows Phone Failed Spectacularly*, HISTORY–COMPUTER (Aug. 8, 2023), https://history-computer.com/the-real-reason-windows-phone-failed-spectacularly.

[73] *Introducing the GPT Store*, OPEN AI (Jan. 10, 2024), https://openai.com/blog/introducing-the-gpt-store.

[74] *See* Michael Schade, *How ChatGPT and Our Language Models Are Developed*, OPENAI, https://help.openai.com/en/articles/7842364-how-chatgpt-and-our-language-models-are-developed; Sreejani Bhattacharyya, *Interesting Innovations from OpenAI in 2021*, AIM (Jan. 1, 2022), https://analyticsindiamag.com/interesting-innovations-from-openai-in-2021; Danny Hernadez & Tom B. Brown, *Measuring the Algorithmic Efficiency of Neural Networks*, ARXIV (May 8, 2020), https://arxiv.org/abs/2005.04305.

[75] *See* Yun, *supra* note 26 at 235 ("Even if data is primarily responsible for a platform's quality improvements, these improvements do not simply materialize with the presence of more data—which differentiates the idea of data-driven network effects from direct network effects. A firm needs to intentionally transform raw, collected data into something that provides analytical insights. This transformation involves costs including those associated with data storage, organization, and analytics, which moves the idea of collecting more data away from a strict network effect to more of a 'data opportunity.'").

D. Five Key Takeaways: Reconceptualizing the Role of Data in Generative-AI Competition

As we explain above, data network effects are not the source of barriers to entry that they are sometimes made out to be. The picture is far more nuanced. Indeed, as economist Andres Lerner demonstrated almost a decade ago (and the assessment is only truer today):

> Although the collection of user data is generally valuable for online providers, *the conclusion that such benefits of user data lead to significant returns to scale and to the entrenchment of dominant online platforms is based on unsupported assumptions*. Although, in theory, control of an "essential" input can lead to the exclusion of rivals, a careful analysis of real-world evidence indicates that such concerns are unwarranted for many online businesses that have been the focus of the "big data" debate.[76]

While data can be an important part of the competitive landscape, incumbents' data advantages are far less pronounced than today's policymakers commonly assume. In that respect, five primary lessons emerge:

1. Data can be (very) valuable, but beyond a certain threshold, those benefits tend to diminish. In other words, having the *most* data is less important than having *enough*;
2. The ability to generate valuable information does not depend on the number of users or the amount of data a platform has previously acquired;
3. The most important datasets are not always proprietary;
4. Technological advances and platforms' engineering decisions affect their ability to generate valuable information, and this effect swamps those that stem from the amount of data they own; and
5. *How* platforms use data is arguably more important than *what* data or *how much* data they own.

These lessons have important ramifications for policy debates over the competitive implications of data in technologically evolving areas.

First, it is not surprising that startups, rather than incumbents, have taken an early lead in generative AI (and in Web 2.0 before it). After all, if data-incumbency advantages are small or even nonexistent, then smaller and more nimble players may have an edge over established tech platforms. This is all the more likely given that, despite significant efforts, the biggest tech platforms were unable to offer compelling generative-AI chatbots and image-generation services before the emergence of ChatGPT, Dall-E, Midjourney, etc.

This suggests that, in a process akin to Clayton Christensen's "innovator's dilemma,"[77] something about the incumbent platforms' existing services and capabilities might have been holding them

---

[76] Lerner, *supra* note 67, at 4-5 (emphasis added).

[77] *See* CLAYTON M. CHRISTENSEN, THE INNOVATOR'S DILEMMA: WHEN NEW TECHNOLOGIES CAUSE GREAT FIRMS TO FAIL (2013).

back in this emerging industry. Of course, this does not necessarily mean that those same services or capabilities could not become an advantage when the generative-AI industry starts addressing issues of monetization and scale.[78] But it does mean that assumptions about a firm's market power based primarily on its possession of data are likely to be off the mark.

Another important implication is that, paradoxically, policymakers' efforts to prevent Web 2.0 platforms from competing freely in generative-AI markets may ultimately backfire and lead to *less,* not more, competition. Indeed, OpenAI is currently acquiring a sizeable lead in generative AI. While competition authorities might like to think that other startups will emerge and thrive in this space, it is important not to confuse those desires with reality. While there currently exists a vibrant AI-startup ecosystem, there is at least a case to be made that significant competition for today's AI leaders will come from incumbent Web 2.0 platforms—although nothing is certain at this stage.

Policymakers should beware not to stifle that competition on the misguided assumption that competitive pressure from large incumbents is somehow less valuable to consumers than that which originates from smaller firms. This is particularly relevant in the context of merger control. An acquisition (or an "acqui-hire") by a "big tech" company does not only, in principle, entail a minor risk to harm competition (it is not a horizontal merger),[79] but could create a stronger competitor to the current market leaders.

Finally, even if there were a competition-related market failure to be addressed in the field of generative AI (which is anything but clear), the remedies under contemplation may do more harm than good. Some of the solutions that have been put forward have highly ambiguous effects on consumer welfare. Scholars have shown that, *e.g.*, mandated data sharing—a solution championed by EU policymakers, among others—may sometimes dampen competition in generative AI.[80] This is also true of legislation like the General Data Protection Regulation (GDPR), which makes it harder for firms to acquire more data about consumers—assuming such data is, indeed, useful to generative-AI services.[81]

---

[78] *See* DAVID J. TEECE, DYNAMIC CAPABILITIES AND STRATEGIC MANAGEMENT: ORGANIZING FOR INNOVATION AND GROWTH (2009).

[79] Antitrust merger enforcement has long assumed that horizontal mergers are more likely to cause problems for consumers than vertical mergers. *See:* Geoffrey A. Manne, Dirk Auer, Brian Albrecht, Eric Fruits, Daniel J. Gilman, & Lazar Radic, *Comments of the International Center for Law and Economics on the FTC & DOJ Draft Merger Guidelines*, (Sep. 18, 2023), https://laweconcenter.org/resources/comments-of-the-international-center-for-law-and-economics-on-the-ftc-doj-draft-merger-guidelines.

[80] *See* Hagiu & Wright, *supra* note 30, at 30 ("We use our dynamic framework to explore how data sharing works: we find that it in-creases consumer surplus when one firm is sufficiently far ahead of the other by making the laggard more competitive, but it decreases consumer surplus when the firms are sufficiently evenly matched by making firms compete less aggressively, which in our model means subsidizing consumers less."); *see also* Lerner, *supra* note 67.

[81] *See, e.g.*, Hagiu & Wright, *id.* ("We also use our model to highlight an unintended consequence of privacy policies. If such policies reduce the rate at which firms can extract useful data from consumers, they will tend to increase the incumbent's competitive advantage, reflecting that the entrant has more scope for new learning and so is affected more by such a policy."); Jian Jia, Ginger Zhe Jin, & Liad Wagman, *The Short-Run Effects of the General Data Protection Regulation on Technology*

In sum, it is a flawed understanding of the economics and practical consequences of large agglomerations of data that has led competition authorities to believe data-incumbency advantages are likely to harm competition in generative AI—or even in the data-intensive Web 2.0 markets that preceded it. Indeed, competition or regulatory intervention to "correct" data barriers and data network and scale effects is liable to do more harm than good.

II. Access restrictions to GPU's

Several competition agencies have expressed their concern about possible "choke points" in AI markets, where a small number of companies in position to exploit existing or emerging bottlenecks across the AI stack" could foreclose competitors' access to key inputs.[82] The Discussion Paper posits that:

> To develop high-performance generative AI models, it is crucial to invest in a sufficient quantity of semiconductor chips tailored for this purpose. Companies such as NVIDIA and Google are prominent developers of these specialized semiconductor chips.
>
> …NVIDIA holds approximately 80% of the GPU market share, and several major companies dominate the cloud service market. Consequently, these companies are considered to have a strong position in the generative AI sector.[83]

Despite its high market share, however, Nvidia faces significant competition from longtime makers of microprocessors like Intel or AMD offer or are developing competing GPU's or processors for AI devices (admittedly, with Intel behind).[84] AMD, for instance, has recently launched a new AI chip, the MI32X that will create competitive pressure for Nvidia.[85] Moreover, in another example of how "big tech" firms can be a pro-competitive force, Alphabet, Amazon, Apple, Meta and Microsoft are investing in research and the ability to produce their own chips.[86] Some of them are currently

---

Venture Investment, 40 MARKETING SCI. 593 (2021) (finding GDPR reduced investment in new and emerging technology firms, particularly in data-related ventures); James Campbell, Avi Goldfarb, & Catherine Tucker, *Privacy Regulation and Market Structure*, 24 J. ECON. & MGMT. STRAT. 47 (2015) ("Consequently, rather than increasing competition, the nature of transaction costs implied by privacy regulation suggests that privacy regulation may be anti-competitive.").

[82] *Joint Statement on Competition in Generative AI*, *supra* note 17.

[83] Discussion Paper, *supra* note 2, at 5.

[84] Sean Hollister, *Intel's Gaudi AI chips are far behind Nvidia and AMD, won't even hit $500M goal*, THE VERGE (Oct. 31, 2024), https://www.theverge.com/2024/10/31/24284860/intel-gaudi-wont-meet-500-million-goal

[85] Kif Leswing, *AMD launches AI chip to rival Nvidia's Blackwell*, CNBC (Oct. 10, 2024), https://www.cnbc.com/2024/10/10/amd-launches-mi325x-ai-chip-to-rival-nvidias-blackwell-.html

[86] See, for instance: Emilia David, *Chip race: Microsoft, Meta, Google, and Nvidia battle it out for AI chip supremacy*, THE VERGE (Updated Oct. 31, 2024), https://www.theverge.com/2024/2/1/24058186/ai-chips-meta-microsoft-google-nvidia/archives/3

manufacturing them. Google, for instance, manufactures its own hardware accelerator, the tensor processing unit (TPU), and considers it of similar performance with NVIDIA GPUs.[87]

Given that GPUs are not the only type of processor unit that can be used for AI —in addition to the aforementioned TPUs there are other computer processors like Central Processing Units (CPUs) that are among the available options as part of an AI computing architecture—[88], Nvidia's market share should probably be adjusted and could be much lower than 80%.

In addition to the above, firms that are currently producing chips that are not suitable for AI purposes may have the ability and incentives to enter the market, should Nvidia try to exploit its market power. Indeed, if Nvidia would, for instance, raise the price of its GPU chips, that would send a signal and create incentives to companies making different types of chips to switch its production to GPU chips, or any other type of chips suitable for AI purposes.[89] This is just not a theoretical conjecture. Nvidia, the alleged dominant in the sector, entered the market focused on graphics-based processing for video game, and then took the opportunity that this approach (parallel processing) represented for AI computing.[90]

Finally, in addition to the fact that computing powers providers like Nvidia may not have the ability to restrict supply to essential AI inputs, there is no reason to believe that they have the *incentives* to do so. Even if they are vertically integrated and produce and distribute AI models and applications, companies like Nvidia would lose important sources of revenue if they decline to provide their chips to downstream firms.

III. Merger Policy and AI. In particular, AI partnerships involving "big tech" firms

Policymakers have expressed particular concern about the anticompetitive potential of deals wherein AI startups obtain funding from incumbent tech firms, even in cases where these strategic partnerships cannot be considered mergers in the antitrust sense (because there is no control exercised by one firm over the other). To date, there is no evidence to support differentiated scrutiny for mergers involving AI firms or, in general, firms working with information technology. The view that so-called "killer acquisitions," for instance, pose a significant competition risk in AI markets is not supported

[87] COMPETITION & MARKETS AUTHORITY, AI FOUNDATION MODELS: INITIAL REPORT (2023), at 13.

[88] Google, *What is a GPU and its role in AI?*, https://cloud.google.com/discover/gpu-for-ai (last visited, Nov. 14, 2024).

[89] Depending on the magnitude and costs of such substitution, these firms could be considered in the same relevant market or as potential competition to Nvidia. See: Jorge Padilla, *The Role of Supply-Side Substitution in the Definition of the Relevant Market in Merger Control*, A Report for DG Enterprise A/4, European Commission (Jun., 2001)

[90] Katie Tarasov, *Nvidia CEO Jensen Huang's big bet on A.I. is paying off as his core technology powers ChatGPT*, CNBC (Mar. 7, 2023), https://www.cnbc.com/2023/03/07/nvidia-grew-from-gaming-to-ai-giant-and-now-powering-chatgpt.html

by solid evidence.[91] To the contrary, there is reason to believe these acquisitions bolster competition by allowing larger firms to acquire capabilities relevant to innovation, and by increasing incentives to invest for startup founders.[92]

Companies with "deep pockets" that invest in AI startups may provide those firms the resources to compete with prevailing market leaders. Firms like Amazon, Google, Meta, and Microsoft, for instance, have been investing to create their own microchips capable of building AI systems, aiming to be less dependent on Nvidia.[93] The tributaries of this flow of funds could serve to enhance competition at all levels of the AI industry.[94]

A. Existing AI Partnerships Are Unlikely to Be Anticompetitive

Some jurisdictions have also raised concerns regarding recent partnerships among big tech firms and AI "unicorns,"[95] in particular, Amazon's partnership with Anthropic; Microsoft's partnership with Mistral AI; and Microsoft's hiring of former Inflection AI employees (including, notably, founder Mustafa Suleyman) and related arrangements with the company. Publicly available information, however, suggests that these transactions may not warrant merger-control investigation, let alone the heightened scrutiny that comes with potential Phase II proceedings. At the very least, given the AI industry's competitive landscape, there is little to suggest these transactions merit closer scrutiny than similar deals in other sectors.

Overenforcement in the field of generative AI could paradoxically engender the very harms that policymakers are seeking to avert. Preventing big tech firms from competing in these markets (for example, by threatening competition intervention as soon as they build strategic relationships with AI startups) may thwart an important source of competition needed to keep today's leading generative-AI firms in check. In short, while competition in AI markets is important,[96] trying naïvely to hold incumbent (in adjacent markets) tech firms back, out of misguided fears they will come to dominate this space, is likely to do more harm than good.

At a more granular level, there are important reasons to believe these kinds of agreements will have no negative impact on competition and may, in fact, benefit consumers—*e.g.*, by enabling those

---

[91] *See* Jonathan M. Barnett, *"Killer Acquisitions" Reexamined: Economic Hyperbole in the Age of Populist Antitrust*, 3 U. CHI. BUS. L. REV. 39 (2023).

[92] *Id.* at 85. ("At the same time, these transactions enhance competitive conditions by supporting the profit expectations that elicit VC investment in the startups that deliver the most transformative types of innovation to the biopharmaceutical ecosystem (and, in some cases, mature into larger firms that can challenge incumbents).)"

[93] Cade Metz, Karen Weise, & Mike Isaac, *Nvidia's Big Tech Rivals Put Their Own A.I. Chips on the Table*, N.Y. TIMES (Jan. 29, 2024), https://www.nytimes.com/2024/01/29/technology/ai-chips-nvidia-amazon-google-microsoft-meta.html.

[94] *See, e.g.*, Chris Metinko, *Nvidia's Big Tech Rivals Put Their Own A.I. Chips on the Table*, CRUNCHBASE (Jun. 12, 2024), https://news.crunchbase.com/ai/msft-nvda-lead-big-tech-startup-investment.

[95] *CMA Seeks Views on AI Partnerships and Other Arrangements*, COMPETITION AND MKTS. AUTH. (Apr. 24, 2024), https://www.gov.uk/government/news/cma-seeks-views-on-ai-partnerships-and-other-arrangements.

[96] As noted *infra*, companies offer myriad "AI" products and services, and specific relevant markets would need to be defined before assessing harm to competition in specific cases.

startups to raise capital and deploy their services at an even larger scale. In other words, they do not bear any of the *prima facie* traits of "killer acquisitions," or even of the acquisition of "nascent potential competitors."[97]

Most importantly, these partnerships all involve the acquisition of minority stakes and do not entail any change of control over the target companies. Amazon, for instance, will not have "ownership control" of Anthropic. The precise amount of shares acquired has not been made public, but a reported investment of $4 billion in a company valued at $18.4 billion does not give Amazon a majority stake or sufficient voting rights to control the company or its competitive strategy. [98] It has also been reported that the deal will not give Amazon any seats on the Anthropic board or special voting rights (such as the power to veto some decisions).[99] There is thus little reason to believe Amazon has acquired indirect or *de facto* control over Anthropic.

Microsoft's investment in Mistral AI is even smaller, in both absolute and relative terms. Microsoft is reportedly investing just $16 million in a company valued at $2.1 billion.[100] This represents less than 1% of Mistral's equity, making it all but impossible for Microsoft to exert any significant control or influence over Mistral AI's competitive strategy. There have similarly been no reports of Microsoft acquiring seats on Mistral AI's board or any special voting rights. We can therefore be confident that the deal will not affect competition in AI markets.

Much the same applies to Microsoft's dealings with Inflection AI. Microsoft hired two of the company's three founders (which currently does not fall under the scope of merger laws), and also paid $620 million for nonexclusive rights to sell access to the Inflection AI model through its Azure Cloud.[101] Admittedly, the latter could entail (depending on deal's specifics) some limited control over Inflection AI's competitive strategy, but there is currently no evidence to suggest this will be the case.

---

[97] *Start-ups, Killer Acquisitions and Merger Control*, OECD (2020), *available at* https://web-archive.oecd.org/2020-10-16/566931-start-ups-killer-acquisitions-and-merger-control-2020.pdf.

[98] Kate Rooney & Hayden Field, *Amazon Spends $2.75 Billion on AI Startup Anthropic in Its Largest Venture Investment Yet*, CNBC (Mar. 27, 2024), https://www.cnbc.com/2024/03/27/amazon-spends-2point7b-on-startup-anthropic-in-largest-venture-investment.html.

[99] *Id.*

[100] Tom Warren, *Microsoft Partners with Mistral in Second AI Deal Beyond OpenAI*, THE VERGE (Feb. 26, 2024), https://www.theverge.com/2024/2/26/24083510/microsoft-mistral-partnership-deal-azure-ai.

[101] Mark Sullivan, *Microsoft's Inflection AI Grab Likely Cost More Than $1 Billion, Says An Insider (Exclusive)*, FAST COMPANY (Mar. 26, 2024), https://www.fastcompany.com/91069182/microsoft-inflection-ai-exclusive; *see also*, *Mustafa Suleyman, DeepMind and Inflection Co-Founder, Joins Microsoft to Lead Copilot*, MICROSOFT CORPORATE BLOGS (Mar. 19, 2024), https://blogs.microsoft.com/blog/2024/03/19/mustafa-suleyman-deepmind-and-inflection-co-founder-joins-microsoft-to-lead-copilot; Krystal Hu & Harshita Mary Varghese, *Microsoft Pays Inflection $ 650 Mln in Licensing Deal While Poaching Top Talent, Source Says*, REUTERS (Mar. 21, 2024), https://www.reuters.com/technology/microsoft-agreed-pay-inflection-650-mln-while-hiring-its-staff-information-2024-03-21; *The New Inflection: An Important Change to How We'll Work*, INFLECTION (Mar. 19, 2024), https://inflection.ai/the-new-inflection; Julie Bort, *Here's How Microsoft Is Providing a 'Good Outcome' for Inflection AI VCs, as Reid Hoffman Promised*, TECH CRUNCH (Mar. 21, 2024), https://techcrunch.com/2024/03/21/microsoft-inflection-ai-investors-reid-hoffman-bill-gates.

Finally, none of these deals entail any competitively significant behavioral commitments from the target companies. There are no reports of exclusivity agreements or other commitments that would restrict third parties' access to these firms' underlying AI models. Again, this means the deals are extremely unlikely to negatively impact the competitive landscape in these markets.

B. AI Partnerships Increase Competition

As discussed in the previous section, the AI partnerships that have recently grabbed antitrust headlines are unlikely to harm competition. They do, however, have significant potential to bolster competition in generative-AI markets by enabling new players to scale up rapidly and to challenge more established players by leveraging the resources of incumbent tech platforms.

The fact that AI startups willingly agree to the aforementioned AI partnerships suggests this source of funding presents unique advantages for them, or they would have pursued capital through other venues. The question for antitrust policymakers is whether this advantage is merely an anticompetitive premium, paid by big tech platforms to secure monopoly rents, or whether the investing firms are bringing something else to the table. As we discussed in the previous section, there is little reason to believe these partnerships are driven by anticompetitive motives. More importantly, however, these deals may present important advantages for AI startups that, in turn, are likely to boost competition in these burgeoning markets.

To start, partnerships with so-called big tech firms are likely a way for AI startups to rapidly obtain equity financing. While this lies beyond our area of expertise, there is ample economic literature to suggest that debt and equity financing are not equivalent for firms.[102] Interestingly for competition

---

[102] *See, e.g.,* Paul Marsh, *The Choice Between Equity and Debt: An Empirical Study*, 37 THE J. OF FINANCE 121, 142 (1982) ("First, it demonstrates that companies are heavily influenced by market conditions and the past history of security prices in choosing between equity and debt. Indeed, these factors appeared to be far more significant in our model than, for example, other variables such as the company's existing financial structure. Second, this study provides evidence that companies do appear to make their choice of financing instrument as though they had target levels in mind for both the long term debt ratio, and the ratio of short term to total debt. Finally, the results are consistent with the notion that these target levels are themselves functions of company size, bankruptcy risk, and asset composition."); *see also,* Armen Hovakimian, Tim Opler, & Sheridan Titman, *The Debt-Equity Choice*, 36 J. OF FINANCIAL AND QUANTITATIVE ANALYSIS 1, 3(2001) ("Our results suggest that, although pecking order considerations affect corporate debt ratios in the short-run, firms tend to make financing choices that move them toward target debt ratios that are consistent with tradeoff models of capital structure choice. For example, our findings confirm that more profitable firms have, on average, lower leverage ratios. But we also find that more profitable firms are more likely to issue debt rather than equity and are more likely to repurchase equity rather than retire debt. Such behavior is consistent with our conjecture that the most profitable firms become under-levered and that firms' financing choices tend to offset these earnings-driven changes in their capital structures."): *see also*, Sabri Boubaker, Wael Rouatbi, & Walid Saffar, *The Role of Multiple Large Shareholders in the Choice of Debt Source*, 46 FINANCIAL MANAGEMENT 241, 267 (2017) ("Our analysis shows that firms controlled by more than one large shareholder tend to rely more heavily on bank debt financing. Moreover, we find that the proportion of bank debt in total debt is significantly higher for firms with higher contestability of the largest controlling owner's power.").

policy, there is evidence to suggest firms tend to favor equity over debt financing when they operate in highly competitive product markets.[103]

Furthermore, there may be reasons that AI startups to turn to incumbent big tech platforms to obtain financing, rather than to other partners (though there is evidence these firms are also raising significant amounts of money from other sources).[104] In short, big tech platforms have a longstanding reputation for deep pockets, as well as a healthy appetite for risk. Because of the relatively small amounts at stake—at least, relative to the platforms' market capitalizations—these firms may be able to move faster than rivals, for whom investments of this sort may present more significant risks. This may be a key advantage in the fast-paced world of generative AI, where obtaining funding and scaling rapidly could be the difference between becoming the next GAFAM or an also-ran.

Partnerships with incumbent tech platforms may also create valuable synergies that enable startups to extract better terms than would otherwise be the case (because the deal creates more surplus for parties to distribute among themselves). Potential synergies include better integrating generative-AI services into existing platforms; several big tech platforms appear to see the inevitable integration of AI into their services as a challenge similar to the shift from desktop to mobile internet, which saw several firms thrive, while others fell by the wayside.[105]

Conversely, incumbent tech platforms may have existing infrastructure that AI startups can use to scale up faster and more cheaply than would otherwise be the case. Running startups' generative-AI services on top of this infrastructure may enable much faster deployment of generative-AI

---

[103] Sabri Boubaker, Walid Saffar, & Syrine Sassi, *Product Market Competition and Debt Choice*, 49 J. OF CORP. FINANCE 204, 208 (2018). ("Our findings that firms substitute away from bank debt when faced with intense market pressure echo the intuition in previous studies that the disciplinary force of competition substitutes for the need to discipline firms through other forms of governance.").

[104] *See, e.g.*, George Hammond, *Andreessen Horowitz Raises $7.2bn and Sets Sights on AI Start-ups*, FINANCIAL TIMES (Apr. 16, 2024), https://www.ft.com/content/fdef2f53-f8f7-4553-866b-1c9bfdbeea42; *Elon Musk's xAI Says It Raised $6 Billion to Develop Artificial Intelligence*, MONEYWATCH (May. 27, 2024), https://www.cbsnews.com/news/elon-musk-xai-6-billion; Krystal Hu, *AI Search Startup Genspark Raises $60 Million in Seed Round to Challenge Google*, REUTERS (Jun. 18, 2024), https://www.reuters.com/technology/artificial-intelligence/ai-search-startup-genspark-raises-60-million-seed-round-challenge-google-2024-06-18; *Visa to Invest $100 Million in Generative AI for Commerce and Payments*, PMYNTS (Oct. 2, 2023), https://www.pymnts.com/artificial-intelligence-2/2023/visa-to-invest-100-million-in-generative-ai-for-commerce-and-payments.

[105] *See, e.g.*, *Eze Vidra, Is Generative AI the Biggest Platform Shift Since Cloud and Mobile?*, VC CAFE (Mar. 6, 2023), https://www.vccafe.com/2023/03/06/is-generative-ai-the-biggest-platform-shift-since-cloud-and-mobile. *See also, OpenAI and Apple Announce Partnership to Integrate ChatGPT into Apple Experiences*, OPENAI (Jun. 10, 2024), https://openai.com/index/openai-and-apple-announce-partnership ("Apple is integrating ChatGPT into experiences within iOS, iPadOS, and macOS, allowing users to access ChatGPT's capabilities—including image and document understanding—without needing to jump between tools."). *See also*, Yusuf Mehdi, *Reinventing Search With a new AI-powered Microsoft Bing and Edge, Your Copilot for the Web*, MICROSOFT OFFICIAL BLOG (Feb. 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 ("'AI will fundamentally change every software category, starting with the largest category of all – search,' said Satya Nadella, Chairman and CEO, Microsoft. 'Today, we're launching Bing and Edge powered by AI copilot and chat, to help people get more from search and the web.'").

technology.[106] Importantly, if these joint strategies entail relationship-specific investments on the part of one or both partners, then big tech platforms taking equity positions in AI startups may be an important facilitator to prevent holdup.[107] Both of these possibilities are perfectly summed up by Swami Sivasubramanian, Amazon's vice president of Data and AI, when commenting on Amazon's partnership with Anthropic:

> Anthropic's visionary work with generative AI, most recently the introduction of its state-of-the art Claude 3 family of models, combined with Amazon's best-in-class infrastructure like AWS Tranium and managed services like Amazon Bedrockfurther unlocks exciting opportunities for customers to quickly, securely, and responsibly innovate with generative AI. Generative AI is poised to be the most transformational technology of our time, and we believe our strategic collaboration with Anthropic will further improve our customers' experiences, and look forward to what's next.[108]

All of this can be expected to have a knock-on effect on innovation and competition in generative-AI markets. To put it simply, a leading firm like OpenAI might welcome the prospect of competition authorities blocking the potential funding of one of its rivals. It may also stand to benefit if incumbent tech firms are prevented from rapidly upping their generative-AI game via partnerships with other AI startups. In short, preventing AI startups from obtaining funding from big tech platforms could not only arrest those startups' growth, but also harm long-term competition in the burgeoning AI industry.

Regarding the "cornering" on specialized talent in AI markets, it should be taken into account that, one of the sources of intense competition in these markets is precisely the will of AI engineers to work more independently and with more agility in their own startups. The Economist, for instance, reports that *"all eight authors of 'Attention is all you need', a paper published in 2017 that provided the algorithmic underpinnings of generative AI, have left Google, where they worked at the time. Seven have founded their own firms (the other joined OpenAI)"*.[109] The same article notes that the supply or AI labor is growing: *"According to a report from Stanford University, in 2011 about 41% of AI PhDs took jobs in industry, roughly the same share as those taking jobs in academia. By 2022 that figure stood at 71% for industry,*

---

[106] *See, e.g.*, *Amazon and Anthropic Deepen Their Shared Commitment to Advancing Generative AI*, AMAZON (Mar. 27, 2024), https://www.aboutamazon.com/news/company-news/amazon-anthropic-ai-investment ("Global organizations of all sizes, across virtually every industry, are already using Amazon Bedrock to build their generative AI applications with Anthropic's Claude AI. They include ADP, Amdocs, Bridgewater Associates, Broadridge, CelcomDigi, Clariant, Cloudera, Dana-Farber Cancer Institute, Degas Ltd., Delta Air Lines, Druva, Enverus, Genesys, Genomics England, GoDaddy, HappyFox, Intuit, KT, LivTech, Lonely Planet, LexisNexis Legal & Professional, M1 Finance, Netsmart, Nexxiot, Parsyl, Perplexity AI, Pfizer, the PGA TOUR, Proto Hologram, Ricoh USA, Rocket Companies, and Siemens.").

[107] Ownership of another firm's assets is widely seen as a solution to contractual incompleteness. *See, e.g.*, Sanford J. Grossman & Oliver D. Hart, *The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration*, 94 J. POLIT. ECON. 691, 716 (1986) ("When it is too costly for one party to specify a long list of the particular rights it desires over another party's assets, then it may be optimal for the first party to purchase all rights except those specifically mentioned in the contract. Ownership is the purchase of these residual rights of control.").

[108] *See* Amazon Staff, *supra* note 106.

[109] *The war for AI talent is heating up*, THE ECONOMIST (Jun. (, 2024), https://www.economist.com/business/2024/06/08/the-war-for-ai-talent-is-heating-up

compared with 20% for academia".[110] This, in turn is creating demand for more AI careers from students, and colleges and universities have all the incentives to meet that demand.[111]

Although, of course, large firms may be able to acquire talent more easily; that is also part of a positive competitive dynamic where employees benefit (with more compensation or better conditions) for the high demand of skilled labor.

Moreover, AI is an industry where engineers and other specialists are concerned not only with income, but also academic prestige. Both independent researchers and specialist working for tech companies, more often than not, publish their research, facilitating the spread of information. Google, for instance, published the first paper on transformer models,[112] and Microsoft published the LoRA paper.[113]

IV. Market Definition in AI

The question of market definition, long a cornerstone of antitrust analysis, is of particular importance and complexity in the context of AI. The difficulty in defining relevant markets accurately stems not only from the novelty of AI technologies, but from their inherent heterogeneity and the myriad ways they intersect with existing markets and business models. In short, it is not yet clear how to determine the boundaries of markets for AI-powered products. Indeed, traditional approaches to market definition will ultimately provide the correct tools to accomplish this task, but, as we discuss below, we do not yet know the right questions to ask.

Regulators and policymakers must develop a nuanced understanding of AI markets, one that moves beyond broad generalizations and marketing hyperbole to examine the specific characteristics of these emerging technologies and their impacts on various product and service markets.

There are three main things that need to be at the forefront of competition authorities' minds when they think about market definition in AI products and services. First, they must understand that AI is not a single thing, but is a composite category composed of many distinct goods and services. Second, and related to looking beyond the AI marketing hype, they must recognize how the extremely heterogeneous products landscape of "AI" intersects with an equally variegated consumer-demand landscape. Finally, they must acknowledge how little we know about these nascent markets,

---

[110] *Id.*

[111] Delilah Brumer and Jeremy Garza, *California students want careers in AI. Here's how colleges are meeting that demand*, A.P. (Oct. 29, 2024), https://apnews.com/us-news/artificial-intelligence-colleges-and-universities-california-nvidia-corp-san-jose-7420d86cc44c9d7b5501af5d2ba94c4a

[112] Jakob Uszkoreit, *Transformer: A Novel Neural Network Architecture for Language Understanding,* Google Research Blog (Aug. 31, 2017), https://research.google/blog/transformer-a-novel-neural-network-architecture-for-language-understanding/

[113] Edward J. Hu, Yelong Shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, Lu Wang & Weizhu Chen, *LoRA: Low-Rank Adaptation of Large Language Models*, Microsoft (Apr. 2022), https://www.microsoft.com/en-us/research/publication/lora-low-rank-adaptation-of-large-language-models/

and that the most important priority at the moment is simply to ask the right questions that will lead to sound competition policy.

A. AI Is Difficult to Define and Not Monolithic

The task of defining AI for the purposes of antitrust analysis is fraught with complexity, stemming from the multifaceted nature of AI technologies and their diverse applications across industries. It is imperative to recognize that AI does not constitute a monolithic entity or a singular market, but rather encompasses a heterogeneous array of technologies, techniques, and applications that defy simplistic categorization.[114]

At its core, the "AI Stack" comprises multiple layers of interrelated yet distinct technological components. At the foundational level, we find specialized hardware such as semiconductors, graphics processing units (GPUs), and tensor processing units (TPUs), as well as other specialized chipsets designed to accelerate the computationally intensive tasks associated with AI. These hardware components, while critical to AI functionality, also serve broader markets beyond AI applications (*e.g.*, crypto and gaming), complicating efforts to delineate clear market boundaries.

The data layer presents another dimension of complexity. AI systems rely on vast quantities of both structured and unstructured data for training and operation.[115] The sourcing, curation, and preparation of this data constitute distinct markets within the AI ecosystem, each with its own competitive dynamics and potential barriers to entry.

Moving up the stack, we encounter the algorithmic layer, where a diverse array of machine-learning techniques—including, but not limited to, supervised learning, unsupervised learning, and reinforcement learning[116]—are employed. These algorithmic approaches, while fundamental to AI functionality, are not uniform in their application or market impact. Different AI applications may utilize

---

[114] As the National Security Commission on Artificial Intelligence has observed:

> AI is not a single technology breakthrough… The race for AI supremacy is not like the space race to the moon. AI is not even comparable to a general-purpose technology like electricity. However, what Thomas Edison said of electricity encapsulates the AI future: "It is a field of fields … it holds the secrets which will reorganize the life of the world." Edison's astounding assessment came from humility. All that he discovered was "very little in comparison with the possibilities that appear."

National Security Commission on Artificial Intelligence, Final Report, 7 (2021), *available at* https://www.dwt.com/-/media/files/blogs/artificial-intelligence-law-advisor/2021/03/nscai-final-report-2021.pdf.

[115] See, *e.g.*, *Structured vs Unstructured Data*, IBM CLOUD EDUCATION (Jun. 29, 2021), https://www.ibm.com/think/topics/structured-vs-unstructured-data; Dongdong Zhang, *et al.*, *Combining Structured and Unstructured Data for Predictive Models: A Deep Learning Approach*, BMC MEDICAL INFORMATICS AND DECISION MAKING (Oct. 29, 2020), https://link.springer.com/article/10.1186/s12911-020-01297-6 (describing generally the use of both structured and unstructured data in predictive models for health care).

[116] For a somewhat technical discussion of all three methods, *see generally* Eric Benhamou, *Similarities Between Policy Gradient Methods (PGM) in Reinforcement Learning (RL) and Supervised Learning (SL)*, SSRN (2019), https://ssrn.com/abstract=3391216.

distinct combinations of these techniques,[117] potentially serving disparate markets and consumer needs.

At the application level, the heterogeneity of AI becomes most apparent. From natural-language processing and computer vision to predictive analytics and autonomous vehicles, AI technologies manifest in a multitude of forms, each potentially constituting a distinct relevant market for antitrust purposes. Moreover, these AI applications can intersect with and compete against non-AI solutions, further blurring the boundaries of what might be considered an "AI market."

The deployment models for AI technologies add yet another layer of complexity to the task of defining antitrust-relevant markets. Cloud-based AI services, edge-computing solutions, and on-premises AI deployments may each serve different market segments and face distinct competitive pressures. The ability of firms to make "build or buy" decisions regarding AI capabilities further complicates the delineation of clear market boundaries.[118]

B. Look Beyond the Marketing Hype

The application of antitrust principles to AI markets necessitates a rigorous analytical approach that transcends superficial categorizations and marketing rhetoric. It is imperative for enforcement authorities to eschew preconceived notions and popular narratives surrounding AI, and to focus instead on empirical evidence and careful economic analysis, in order to accurately assess competitive dynamics in AI-adjacent markets.

The allure of AI as a revolutionary technology has led to a proliferation of marketing claims and industry hype[119] that often may obscure the true nature and capabilities of AI systems. This obfuscation presents a significant challenge for antitrust authorities, who must disentangle factual competitive realities from speculative or exaggerated assertions about AI's market impact. This task is further complicated by the rapid pace of technological advancement in the field, which can render even recent market analyses obsolete.

A particularly pernicious misconception that must be addressed is the notion that AI technologies operate in a competitive vacuum, distinct from and impervious to competition from non-AI alternatives. This perspective risks leading antitrust authorities to define markets too narrowly, potentially

---

[117] *Id.*

[118] For a discussion of the "buy vs build" decisions firms employing AI undertake, *see* Jonathan M. Barnett, *The Case Against Preemptive Antitrust in the Generative Artificial Intelligence Ecosystem*, in ARTIFICIAL INTELLIGENCE AND COMPETITION POLICY (Alden Abbott and Thibault Schrepel eds., 2024), at 3-6.

[119] See, *e.g.*, Melissa Heikkilä & Will Douglas Heaven, *What's Next for AI in 2024*, MIT TECH. REV. (Jan. 4, 2024), https://www.technologyreview.com/2024/01/04/1086046/whats-next-for-ai-in-2024 (Runway hyping Gen-2 as a major film-production tool that, to date, still demonstrates serious limitations). LLMs, impressive as they are, have been touted as impending replacements for humans across many job categories, but still demonstrate many serious limitations that may ultimately limit their use cases. *See, e.g.*, Melissa Malec, *Large Language Models: Capabilities, Advancements, And Limitations*, HATCHWORKSAI (Jun. 14, 2024), https://hatchworks.com/blog/gen-ai/large-language-models-guide.

overlooking significant competitive constraints from traditional technologies or human-driven services.

Consider, for instance, the domain of natural-language processing. While AI-powered language models have made significant strides in recent years, they often compete directly with human translators, content creators, and customer-service representatives. Similarly, in the realm of data analysis, AI systems may vie for market share not only with other AI solutions, but also with traditional statistical methods and human analysts. Failing to account for these non-AI competitors in market-definition exercises could result in a distorted view of market power and competitive dynamics.

Moreover, the tendency to treat AI as a monolithic entity obscures the reality that many AI-powered products and services are, in fact, hybrid solutions that combine AI components with traditional software and human oversight.[120] This hybridization further complicates market-definition efforts, as it becomes necessary to assess the degree to which the AI element of a product or service contributes to its market position and substitutability.

C. Current Lack of Knowledge About Relevant Markets

It is crucial to acknowledge at this juncture the profound limitations in our current understanding of how AI technologies will ultimately shape competitive landscapes across various industries. This recognition of our informational constraints should inform a cautious and empirically grounded approach to market definition in the context of AI.

The dynamic nature of AI development renders many traditional metrics for market definition potentially unreliable or prematurely restrictive. Market share, often a cornerstone of antitrust analysis, may prove particularly volatile in AI markets, where technological breakthroughs can rapidly alter competitive positions. Moreover, the boundaries between distinct AI applications and markets remain fluid, with innovations in one domain frequently finding unexpected applications in others, and thereby further complicating efforts to delineate stable market boundaries.

In this context, Jonathan Barnett's observations regarding the dangers of preemptive antitrust approaches in nascent markets are particularly salient.[121] Barnett argues persuasively that, at the early stages of a market's development, uncertainty concerning the competitive effects of certain business

---

[120] *See, e.g.*, *Hybrid AI: A Comprehensive Guide to Applications and Use Cases*, SOLULAB, https://www.solulab.com/hybrid-ai (last visited Jul. 12, 2024); *Why Hybrid Intelligence Is the Future of Artificial Intelligence at McKinsey*, MCKINSEY & CO. (Apr. 29, 2022), https://www.mckinsey.com/about-us/new-at-mckinsey-blog/hybrid-intelligence-the-future-of-artificial-intelligence; Vahe Andonians, *Harnessing Hybrid Intelligence: Balancing AI Models and Human Expertise for Optimal Performance*, COGNAIZE (Apr. 11, 2023), https://blog.cognaize.com/harnessing-hybrid-intelligence-balancing-ai-models-and-human-expertise-for-optimal-performance; *Salesforce Artificial Intelligence*, SALESFORCE, https://www.salesforce.com/artificial-intelligence (last visited Jul. 12, 2024) (combines traditional CRM and algorithms with AI modules); *AI Overview*, ADOBE, https://www.adobe.com/ai/overview.html (last visited Jul. 12, 2024) (Adobe packages generative AI tools into its general graphic-design tools).

[121] Barnett *supra* note 118.

practices is likely to be especially high.[122] This uncertainty engenders a significant risk of false-positive error costs, whereby preemptive intervention may inadvertently suppress practices that are either competitively neutral or potentially procompetitive.[123]

The risk of regulatory overreach is particularly acute in the realm of AI, where the full spectrum of potential applications and competitive dynamics remains largely speculative. Premature market definition and subsequent enforcement actions based on such definitions could stifle innovation and impede the natural evolution of AI technologies and business models.

Further complicating matters is the fact that what constitutes a relevant product in AI markets is often ambiguous and subject to rapid change. The modular nature of many AI systems, where components can be combined and reconfigured to serve diverse functions, challenges traditional notions of product markets. For instance, a foundational language model might serve as a critical input for a wide array of downstream applications, from chatbots to content-generation tools, each potentially constituting a distinct product market. The boundaries between these markets, and the extent to which they overlap or remain distinct, are likely to remain in flux in the near future.

Given these uncertainties, antitrust authorities must adopt a posture of epistemic humility when approaching market definition in the context of AI. This approach of acknowledged uncertainty and adaptive analysis does not imply regulatory paralysis. Rather, it calls for a more nuanced and dynamic form of antitrust oversight, one that remains vigilant to potential competitive harms while avoiding premature or overly rigid market definitions that could impede innovation.

Market definition should reflect our best understanding of both AI and AI markets. Since this understanding is still very much in an incipient phase, antitrust authorities should view their current efforts not as definitive pronouncements on the structure of AI markets, but as iterative steps in an ongoing process of learning and adaptation. By maintaining this perspective, regulators can hope to strike a balance between addressing legitimate competitive concerns and fostering an environment conducive to continued innovation and dynamic competition in the AI sector.

D. Key Questions to Ask

Finally, the most important function for enforcement authorities to play at the moment is to ask the right questions that will help to optimally develop an analytical framework of relevant markets in subsequent competition analyses. This framework should be predicated on a series of inquiries designed to elucidate the true nature of competitive dynamics in AI-adjacent markets. While the specific contours of relevant markets may remain elusive, the process of rigorous questioning can provide valuable insights and guide enforcement decisions.

---

[122] *Id.* at 7-8.

[123] *Id.*

Two fundamental questions emerge as critical starting points for any attempt to define relevant markets in AI contexts.

First, "Who are the consumers, and what is the product or service?" This seemingly straightforward inquiry belies a complex web of considerations in AI markets. The consumers of AI technologies and services are often not end-users, but rather, intermediaries that participate in complex value chains. For instance, the market for AI chips encompasses not only direct purchasers like cloud-service providers, but also downstream consumers of AI-powered applications. Similarly, the product or service in question may not be a discrete AI technology, but rather a bundle of AI and non-AI components, or even a service powered by AI but indistinguishable to the end user from non-AI alternatives.

The heterogeneity of AI consumers and products necessitates a granular approach to market definition. Antitrust authorities must carefully delineate between different levels of the AI value chain, considering the distinct competitive dynamics at each level. This may involve separate analyses for markets in AI inputs (such as specialized hardware or training data), AI development tools, and AI-powered end-user applications.

Second, and perhaps more crucially, "Does AI fundamentally transform the product or service in a way that creates a distinct market?" This question is at the heart of the challenge in defining AI markets. It requires a nuanced assessment of the degree to which AI capabilities alter the nature of a product or service from the perspective of consumers.

In some cases, AI's integration into products or services may represent merely an incremental improvement, not warranting the delineation of a separate market. For example, AI-enhanced spell-checking in word-processing software might not constitute a distinct market from traditional spell-checkers if consumers do not perceive a significant functional difference.

Conversely, in other cases, AI may enable entirely new functionalities or levels of performance that create distinct markets. Large language models capable of generating human-like text, for instance, might be considered to operate in a market separate from traditional writing aids or information-retrieval tools (or not, depending on the total costs and benefits of the option).

The analysis must also consider the potential for AI to blur the boundaries between previously distinct markets. As AI systems become more versatile, they may compete across multiple traditional product categories, challenging conventional market definitions.

In addressing these questions, antitrust authorities should consider several additional factors:

1. The degree of substitutability between AI and non-AI solutions, from the perspective of both direct purchasers and end-users.

2. The extent to which AI capabilities are perceived as essential or differentiating factors by consumers in the relevant market.

3. The potential for rapid evolution in AI capabilities and consumer preferences, which may necessitate dynamic market definitions.

4. The presence of switching costs or lock-in effects, which could influence market boundaries.

5. The geographic scope of AI markets, which may transcend traditional national or regional boundaries.

It is crucial to note that these questions do not yield simple or static answers. Rather, they serve as analytical tools to guide ongoing assessment of AI markets. Antitrust authorities must be prepared to revisit and refine their market definitions as technological capabilities evolve and market dynamics shift.

Moreover, the process of defining relevant markets in the context of AI should not be viewed as an end in itself, but as a means to understand competitive dynamics and to inform enforcement decisions. In some cases, traditional market-definition exercises may prove insufficient, necessitating alternative analytical approaches that focus on competitive effects or innovation harms.

By embracing this questioning approach, antitrust authorities can develop a more nuanced and adaptable framework for market definition in AI contexts. This approach would acknowledge the complexities and uncertainties inherent in AI markets, while providing a structured methodology to assess competitive dynamics. As our understanding of AI markets deepens, this framework will need to evolve further, ensuring that antitrust enforcement remains responsive to the unique challenges posed by artificial-intelligence technologies.