Contested Ground: Early Competition and Market Dynamics in Generative AI*

Andrea Asoni† Chiara Farronato‡ Matteo Foschi† Oliver Latham†

June 2025

Abstract

This short paper draws on recent industry developments and an original survey of business leaders to examine whether GenAI markets will remain competitive. It offers insights for business leaders navigating the opportunities offered by GenAI and for policymakers aiming to understand potential risks of market concentration.

The release of ChatGPT in November 2022 marked a turning point for generative artificial intelligence (GenAI). With a hundred million users in just two months, it became the fastest-growing consumer application in history and signaled the arrival of GenAI as a force reshaping industries.[1] What began as a technological breakthrough has quickly become a strategic priority for companies and investors.

*The authors are immensely grateful to Connor Haskin and Hayden Schrauff for invaluable research assistance. As of June 2025, Farronato serves as a member of the Anthropic Economic Advisory Council. This research and the data therein were collected prior to this role. Asoni, Foschi, and Latham have advised multiple technology firms on antitrust issues in the AI space, including Microsoft in respect of investigations into its partnership with OpenAI and its hiring of staff from the start-up Inflection. All views expressed in this paper are those of the authors, not their respective institutions. Portions of this manuscript were assisted by generative AI tools to improve clarity and style. All content was reviewed and edited by the authors.

†Charles River Associates

‡Harvard University, CEPR, NBER

[1] https://www.reuters.com/technology/chatgpt-sets-record-fastest-growing-user-base-analyst-note-2023-02-01/. We note that this record was subsequently broken by Meta's "Threads" app. https://www.itp.net/edge/digital-culture/threads-reaches-100-million-users-dethrones-chatgpt, accessed May 2025.

This paper addresses three central questions. First, is GenAI following the same path as Web 2.0, where early advantages led to long-term dominance by a handful of players? Second, what do supply dynamics, especially current trends on partnerships and deals across the AI value chain, tell us about the likely future industry structure? And third, what do early demand patterns suggest about adoption of GenAI inside organizations? Our observations can inform both policy makers and business leaders interested in expanding and investing in AI.

To explore these questions, we draw from the economic analysis of digital platforms to identify whether the key drivers of market dominance of the Web 2.0 era will play a similar role in GenAI. Further, we examine data on acquisitions along the AI value chain and present results from a survey of business leaders we conducted to understand how GenAI is being adopted inside organizations.

What we find is both encouraging and cautionary. First, the economics of GenAI differ meaningfully from Web 2.0: the powerful data feedback loops and zero-price dynamics that have driven antitrust concerns in search and social media seem to be largely absent; second, there is evidence of competitors, including independent start-ups, investing to catch up with first movers. Third, upstream layers of the AI stack are more concentrated than downstream layers. One needs to consider whether vertical integration within the stack is motivated by efficiencies or a desire to control access to critical inputs. Finally, enterprise users, likely wary of lock-in risks from past digital transitions, are adopting a multihoming approach, engaging with multiple providers simultaneously.

The paper is organized as follows. Section 1 outlines the competition concerns occupying regulators and how these are shaped by key learnings from the evolution of Web 2.0. We describe early market dynamics and industry features in foundation models to evaluate whether they suggest a repeat of the Web 2.0 experience. Section 2 focuses on the supply of GenAI solutions, by investigating the extent of vertical integration in the AI "stack" and its potential effects. Section 3 discusses demand trends and presents survey data on usage of GenAI within organizations. We conclude in Section 4.

1 Is GenAI a Repeat of Web 2.0?

GenAI is expected to revolutionize technology, economy, and society. It is set to become a fundamental component of our day-to-day lives to the point that some have compared it to the discovery of electricity or even fire.[2]

[2]For example, as early as 2018, Google's CEO, Sundar Pichai stated that "AI is probably the most important thing humanity has ever worked on. I think of it as something more

Global regulators have begun investigations into GenAI much earlier than in other industries. A likely motivation is that regulators are aware of their failure to identify (or prevent) concentration trends in various segments during the Web 2.0 era[3] and are now facing immense difficulties in their attempts at ex-post corrections.[4] Furthermore, several of the firms that emerged during the Web 2.0 era are also participating in the GenAI space, which has likely piqued the interest and attention of the antitrust enforcers. Regulators are therefore keen to understand the technology early on so that they can monitor and prevent markets from "tipping" to a small list of large companies.

The debate on the successes and failures of antitrust in the Web 2.0 era has identified certain economic fundamentals that can contribute to market tipping. The first is network effects (Rochet and Tirole, 2006; Parker and Van Alstyne, 2005). Network effects and tipping had originally been identified in the literature much earlier. For example, see the classic work of Katz and Shapiro (1985) Katz and Shapiro (1985). While these had been identified prior to Web 2.0, the major digital platforms were able to rely on both direct and indirect network effects (or "flywheels") to reach unprecedented scale.

Second, data feedback loops played a critical role during the Web 2.0 era: because more users generate better data, which leads to better services, incumbents can become increasingly difficult to dislodge, even if a competitor offers a better product (Hagiu and Wright, 2023).

Third, the prevalence of zero prices is notable: search, social networks, and other Web 2.0 services are often offered at zero pecuniary price because they feature minuscule marginal costs and can be funded via advertising (Bergemann and Bonatti, 2019).

When zero prices, network effects, and data feedback loops converge, they can create a substantial competitive advantage: smaller firms and new entrants often lack the scale to match quality and cannot undercut incumbents on price.

A key question for regulators is whether this tipping dynamic could reoccur in the GenAI space and whether the incumbent players from the Web 2.0 era will be able to leverage

profound than electricity or fire". See https://money.cnn.com/2018/01/24/technology/sundar-pichai-google-ai-artificial-intelligence/index.html. More recently, Jamie Dimon, JPMorgan's CEO stated how "the consequences [of AI] will be extraordinary and possibly as transformational as some of the major technological inventions of the past several hundred years. Think the printing press, the steam engine, electricity, computing and the Internet, among others" - see https://www.nbcnewyork.com/news/business/money-report/jamie-dimon-says-ai-could-be-as-transformative-as-electricity-or-the-internet-heres-how-to-invest/5303274/.

[3]Web 2.0 is generally defined as the period between 2000 and 2010 characterized by a shift from static websites to dynamic, user-driven platforms. See one of the seminal contributions in defining Web 2.0 by O'Reilly (2007).

[4]See, for example, the current FTC lawsuit against Meta challenging their acquisitions of Instagram in 2012 and WhatsApp in 2014, or the DOJ lawsuits against Google for monopolizing digital advertising and search.

their existing assets (e.g. inputs to GenAI technology or distribution channels) to limit competition in GenAI or entrench their positions in their respective industries.

Regulators are actively looking at these questions. The UK Competition and Markets Authority started various market studies on foundation models and partnerships in the GenAI space;[5] the Federal Trade Commission issued a Staff report on partnerships across the GenAI stack;[6] the European Commission released a policy brief on GenAI and potential competition concerns;[7] and multiple national agencies, including in France,[8] Canada,[9] and Portugal[10] have published reports identifying potential concerns.

The situation in 2023 was concerning for people worried about a re-run of the Web 2.0 era. Once again, a new name (OpenAI) had developed a technological edge with the release of GPT-3.5. OpenAI appeared to be on the verge of running away with it, pulling ahead of competitors to the point of becoming uncatchable and potentially entrenching a dominant position for the long run. It was at this point that many of the first regulatory reports on competition in GenAI began to emerge.[11]

The competitive response from existing and new companies proved, however, quite formidable. The UK Competition and Market Authority reported in April 2024 how "the number of [foundation models] globally continues to increase." Over a hundred and twenty models were released between the first report in September 2023 and its update in April 2024. The major public releases are summarized in Figure 1 (from the Competition and Market Authority's update report).[12]

Importantly, these competitors appear to be replicating, and in some cases surpassing, OpenAI's technological capabilities. This is illustrated in Figure 2, which plots the scores of different models on the "Massive Multitask Language Understanding" benchmark.[13] The

[5] See https://www.gov.uk/cma-cases/ai-foundation-models-initial-review.

[6] See https://www.ftc.gov/news-events/news/press-releases/2025/01/ftc-issues-staff-report-ai-partnerships-investments-study.

[7] See https://competition-policy.ec.europa.eu/document/download/c86d461f-062e-4dde-a662-15228d6ca385_en.

[8] See https://www.autoritedelaconcurrence.fr/en/press-release/generative-artificial-intelligence-autorite-issues-its-opinion-competitive.

[9] See https://competition-bureau.canada.ca/en/how-we-foster-competition/education-and-outreach/artificial-intelligence-and-competition.

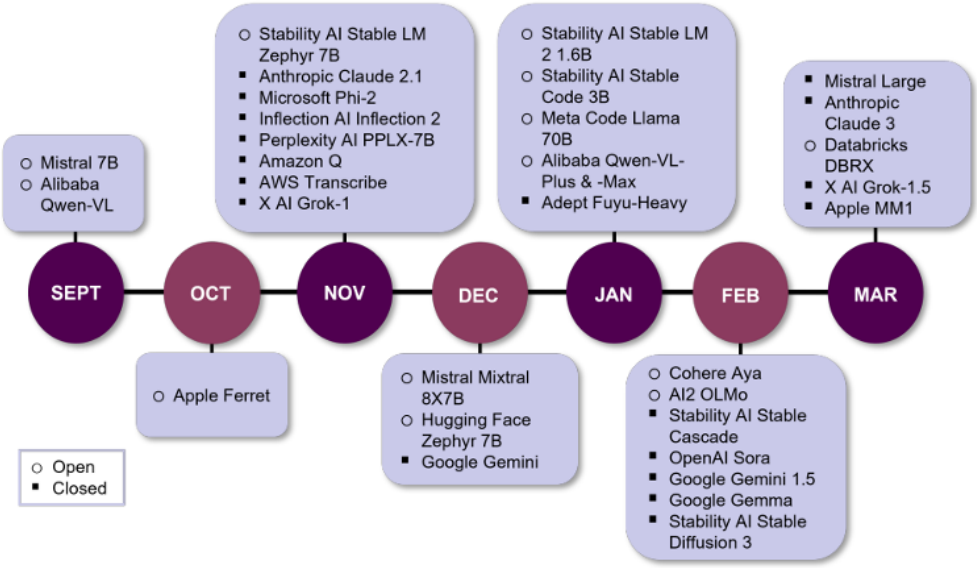
[10] See https://www.concorrencia.pt/en/articles/adc-warns-competition-risks-generative-artificial-intelligence-sector.

[11] For example, the UK Competition and Markets Authority (CMA) launched its initial review of the Foundation Model space in May 2023 to better understand how competitive dynamics were unfolding and whether early advantages could lead to long-term market power.

[12] See UK Competition and Markets Authority's "AI Foundation Models: Technical Update Report" at https://assets.publishing.service.gov.uk/media/661e5a4c7469198185bd3d62/AI_Foundation_Models_technical_update_report.pdf, paragraphs 2.4 and 2.5.

[13] The MMLU is a benchmark that was widely used to assess model's multitasking capabilities across diverse

Figure 1: Publicly Announced Foundation Models Released between September 2023 and March 2024



Source: Competition and Market Authority "AI Foundation Models: Technical Update Report," Figure 2 (Available at https://assets.publishing.service.gov.uk/media/661e5a4c7469198185bd3d62/AI_Foundation_Models_technical_update_report.pdf). The figure shows the release times of the major foundation models. Each model is identified by the organization releasing the model, the model name, and the model version, wherever applicable. The period covered is from September 2023 to March 2024. Only models with official public releases are included.

plot shows that, by the fall of 2024, many foundation models had caught up and surpassed GPT-4's capabilities. At the same time, many other GenAI companies – including new entrants such as Anthropic, Perplexity AI, and Mistral, as well as incumbents like Google, Microsoft, and others – had launched their own GenAI applications, while companies like Adobe developed foundation models for internal integration into their services.[14] More recently, the industry faced another shock with the entry of DeepSeek, a Chinese-based start-up that, in January 2025, released a foundation model allegedly as capable as GPT-4 but trained for a tenth of the cost. On release day, DeepSeek caused US tech stocks to lose almost $1 trillion in market value.[15]

It is worth noting that this is not the first instance of a market leader in GenAI (OpenAI in this case) being matched or overtaken by others. In the 2010s, Google would have been seen as the clear technological leader: it owned DeepMind, the startup responsible for combining deep learning theory with reinforcement learning;[16] it was vertically integrated into cloud computing and downstream applications collecting data; and it contributed to the invention of transformers which laid out the basis for large language models (Vaswani et al., 2017).

Overall, the strength of the competitive response appears inconsistent with the idea that foundation models are heading towards tipping. Yet, the durability of this competitive landscape remains uncertain.

1.1 What Comes Next?

The developments above indicate strong competition in the foundation models' space. Yet an important question remains: is this activity just a temporary surge that will ultimately give way to dominance by a few players? Are we witnessing another Yahoo vs. Google or MySpace vs. Facebook? To answer this question, it is useful to analyze GenAI with respect to the three key drivers of consolidation in the Web 2.0 era that we discussed above: network effects, data feedback loops, and zero prices.[17]

subject including humanities, STEM, and others. Through 2024, the MMLU benchmark remained one of the most widely used tools to evaluate new models. Over time, this benchmark has become "saturated" (in the sense that almost all current models achieve high scores) and attention has shifted to different measures of success – such as MMLU-Pro, a more demand version of the MMLU benchmark.

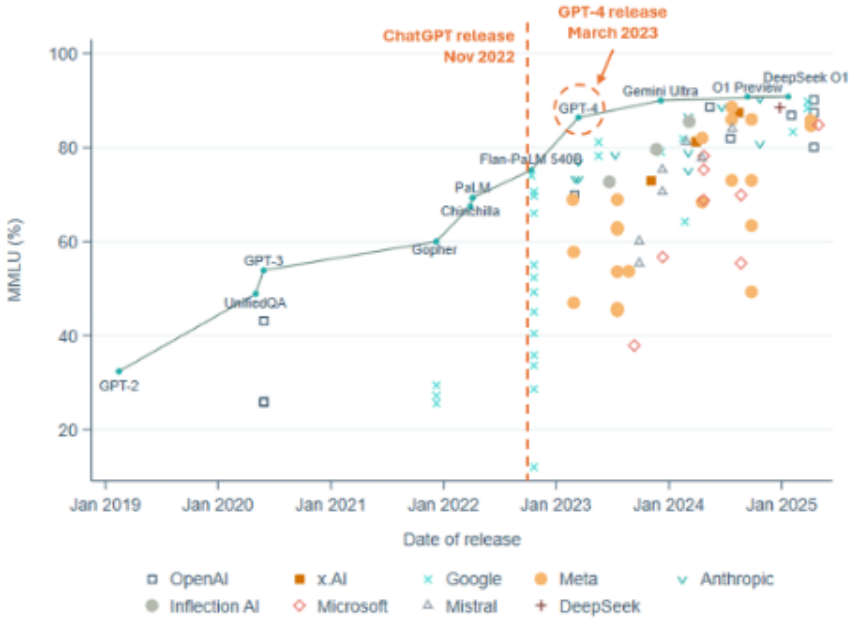
[14] See https://news.adobe.com/news/2025/04/adobe-revolutionizes-ai-assisted-creativity-firefly.

[15] See https://www.investors.com/etfs-and-funds/sectors/sp500-deepseek-ai-sparks-trillion-in-u-s-tech-destruction/.

[16] Deep Learning is recognized as a pivotal change in the evolution of AI – see for example https://epoch.ai/data/notable-ai-models.

[17] For a related discussion of the risk of tipping in foundation model focusing in particular on the role of data feedback loops, see Hagiu and Wright (2025).

Figure 2: Massive Multitask Language Understanding (MMLU) Benchmark Results over Time



Source: Papers with Code, "Multi-task Language Understanding on MMLU", HuggingFace, "LMSYS Chatbot Arena Leaderboard", official release blogs of Microsoft, Adept AI, Inflection, and Anthropic. Data as of April 2025.

First, GenAI is typically not used to facilitate meaningful interactions between users. As a result, the platform flywheel mechanism, where user growth reinforces value, has not taken hold to date. Most GenAI tools today are single-sided, serving individual users rather than facilitating interactions between distinct user groups.

Second, GenAI is currently not exhibiting the same level of data feedback loops we saw in the Web 2.0 era. Models do interact with users, but the complexity of that interaction, compared to search for example, makes it challenging to design a self-improving algorithm of the scale and efficiency of search models. Instead, there may even be risks of model collapse, whereby recursive training on synthetic data influenced by previous GenAI models can lead to negative feedback loops where errors and biases are amplified over successive generations (Shumailov et al., 2023).

Finally, GenAI models are not priced at zero. Inference requires substantial compute and energy resources, and the more capable models tend to be more expensive to query. Most providers offer tiered pricing, with higher priced versions providing more advanced models.[18] This cost structure leads to a more conventional price-quality trade-off, allowing

[18]OpenAI's top tier subscription costs $200 a month, Microsoft charges companies a per-user license for Copilot Pro, Claude 3 Opus is charged at $20 a month. Sam Altman has referred to future models costing $2,000, albeit under an unclear time period. https://www.inc.com/ben-sherry/openais-next-

new entrants to challenge incumbents on price, quality, or both.

Of course, this assessment is based on the current state of GenAI models, which are primarily designed as personal tools. Most GenAI applications today – whether for content generation, coding assistance, or productivity – are geared towards individual use and do not yet exhibit the structural characteristics that supported consolidation in Web 2.0: user-to-user interactions are limited (hindering network effects) and data advantages seem modest (limiting feedback loops). However, it is not a given that this structure will persist. As applications evolve and increasingly rely on proprietary data to fine-tune or personalize offerings, new opportunities may arise to embed GenAI models into traditional two-sided platforms that collect sizable amounts of user data. This, in turn, could revive, or even strengthen, the economic mechanisms that historically led to market tipping in Web 2.0.[19]

At the same time, there are strong signals that the current phase remains open to competition, or that the value created by GenAI is large enough to support multiple players. Indeed, capital markets appear very willing to back new entrants. For example, Elon Musk's xAI and start-ups founded by former OpenAI executives Mira Murati and Ilya Sutskever are reportedly raising capital at valuations in the tens of billions of dollars.[20] This suggests that the window for new entrants to compete remains open, at least for the foreseeable future.

2 Vertical Integration along the AI Stack

Some of the concerns stemming from the Web 2.0 experience are related to how large firms can leverage their position in one sector of the economy to increase control over adjacent segments. Concerns have been expressed about platforms moving up and down the supply chain and the potential anticompetitive effects of these movements.

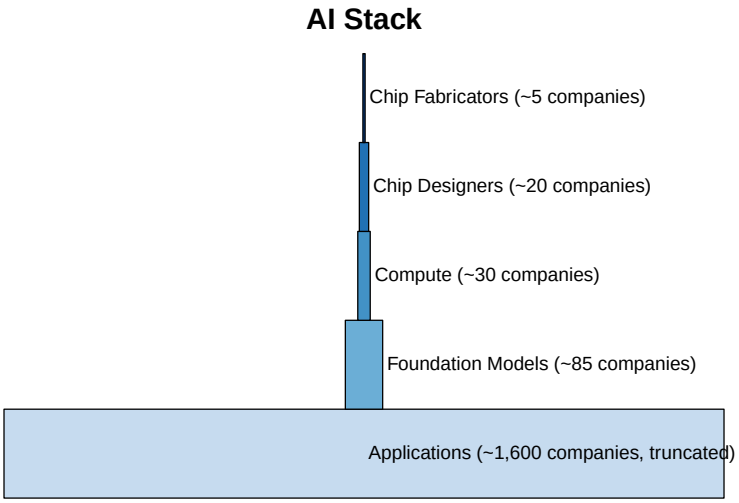
While much of the competition in GenAI discussed above has focused on foundation models and applications, a complete understanding requires looking further upstream in the AI supply chain, often referred to as "AI stack." The development and deployment of GenAI depend on a value chain that includes compute infrastructure, chip design and manufacturing, and that is increasingly becoming a focus of strategic activity. In this section, we examine the extent and implications of vertical integration across the AI stack and consider whether

generation-models-could-reportedly-cost-2000.html.

[19] For example, user-generated fine-tuning data and Reinforcement Learning from Human Feedback could recreate two-sided platform dynamics.

[20] https://techcrunch.com/2025/04/12/openai-co-founder-ilya-sutskevers-safe-superintelligence-reportedly-valued-at-32b/; and https://techcrunch.com/2025/04/10/mira-muratis-ai-startup-is-reportedly-aiming-for-a-massive-2b-seed-round/; and https://techcrunch.com/2025/04/25/musks-xai-holdings-is-reportedly-raising-the-second-largest-private-funding-round-ever/.

Figure 3: Number of Companies across each Layer of the AI Stack



Source: Authors' analysis, based on data from EpochAI and Pitchbook. The list of companies is provided in Online Appendix A.

control over these upstream layers could shape the competitive landscape going forward.

To do this, we collect data on companies active in the AI stack from various sources.[21] We define the AI stack as comprising five layers: chip fabrication, chip design, compute infrastructure, foundation model development, and applications.[22] Figure 3 represents the AI stack as an inverted funnel, showing the number of companies active in each layer. Each segment is proportional to the number of companies, except the applications layer, which is truncated.

The current landscape shows a clear pattern: the number of active firms increases as one moves down the AI stack, with the highest number of firms operating at the application layer

[21]Data on the compute and application layers is from Pitchbook. See https://pitchbook.com/ (accessed May 2025). The criterion to select computing firms is "ai neocloud" (keyword) AND "cloud" (keyword) AND "SaaS" (vertical). The criterion to select applications is {"Artificial Intelligence & Machine Learning" (vertical) AND "artificial intelligence" (keyword) AND "platform" (keyword) AND "developer" OR "Artificial Intelligence & Machine Learning" (vertical) AND "artificial intelligence" (keyword) AND "software" (keyword) AND "developer" (keyword) AND $25 million or more raised or in annual revenue}. This latter data was further filtered to exclude companies that are currently out of business. Data on LLMs and ML hardware is based on Epoch AI, https://epoch.ai/data# accessed April 2025.

[22]Consistent with this, a recent event by the startup incubator YCombinator characterized the AI stack as follows: "your start up is just a GPT wrapper, OpenAI is a Nvidia wrapper, Nvidia is a TSMC wrapper, TSMC is an ASML wrapper, ASML is a Zeiss wrapper, Zeiss is a sand wrapper."

and only a few players (Intel, Samsung, SMIC, TSMC) active in chip fabrication. Setting aside geopolitical complexities, from a regulatory perspective key questions arise: (i) is there evidence of consolidation; (ii) if so, is it primarily vertical (across layers) or horizontal (within layers); and (iii) is consolidation likely to benefit or harm competition and consumers? To explore these questions, we source data from Crunchbase and SDC Platinum on acquisitions involving firms that operate across the GenAI stack between 2015 and 2025.[23]

Figure 4 shows a mix of upstream and downstream vertical integration and strategic investments from players across the stack. Two findings stand out. First, the number of acquisitions completed by Google, Amazon, Apple, Meta, Microsoft, and NVIDIA (GAMMAN)[24]—eighty eight in total—is comparable to the combined number of acquisitions by all other firms operating in the chip (fabrication and design), compute, and model layers, which stands at a hundred and two. Second, most acquisitions are vertical in nature and concentrated in the downstream applications layer.

For example, NVIDIA has expanded downstream from chip design into compute, and more recently into the model and applications layers through acquisitions such as Run:AI (a model orchestration start-up). Although not included in Figure 4, NVIDIA has also pursued partnerships like its collaboration with CoreWeave, a cloud compute provider specialized in AI workloads. Similarly, Microsoft, Amazon, and Google, who have already been active across multiple layers, have continued expanding. Indeed, these major tech firms have started working on chip design (mostly for internal use) and entered strategic partnerships with various model providers, including OpenAI and Anthropic.

Importantly, vertical integration is not limited to the GAMMAN players. Other firms are also pursuing cross-layer integration. Companies like CoreWeave (which acquired Weights Biases and Conductor technologies), Cerebras (which signed strategic partnerships with G42), and others have made strategic acquisitions or forged exclusive partnerships to secure upstream and downstream capabilities.

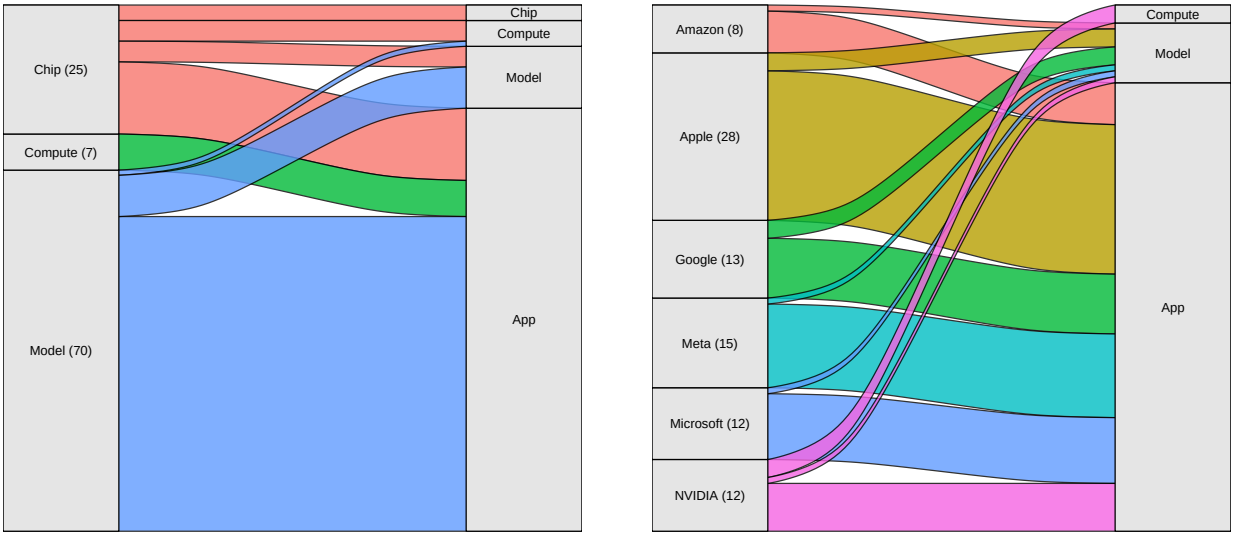
While the flurry of activity across the layers of the AI stack by large and small firms alike is clear, its implications for consumers, potential entrants, and the economy more broadly can have multiple interpretations. A skeptical view holds that these dynamics may ultimately strengthen incumbent firms, allowing them to solidify their existing advantage and leaving limited space for new entrants or smaller competitors.

On the other hand, there are multiple positive implications from the current activity.

[23] See https://www.crunchbase.com/.

[24] GAMMAN is a term coined by the Competition and Market Authority in its market study on AI foundation models which add Nvidia. See https://www.gov.uk/cma-cases/ai-foundation-models-initial-review.

Figure 4: Acquisitions along the AI Stack



(a) Non GAMMAN

(b) GAMMAN

Source: Authors' analysis, based on data from Crunchbase and SDC Platinum. The list of acquisitions is provided in Appendix B.

First, in the context of a major technological shift, common ownership and strategic partnerships can significantly enhance innovation and growth. Because GenAI is emerging as a general-purpose technology with applications across diverse sectors, firms that develop enabling technologies often face under-incentives to innovate when the benefits of their breakthroughs spill over to other players, including competitors (Antón et al., 2025). By forging partnerships or acquiring stakes in adjacent layers of the value chain, firms can partially internalize these externalities. This, in turn, raises the marginal return on investment and strengthens incentives to innovate.

Second, strategic partnerships between firms operating at different layers of the stack can be more cost-efficient than traditional financing methods. For instance, a model developer that secures compute resources at cost through a partnership with a cloud provider can access greater computational capacity than if it raised the same amount of capital from a venture investor and paid market rates. Hence, these arrangements may also reduce the costs of innovation.

Third, strategic partnerships, acquisitions, and expansions can reflect mutual efforts by firms to reduce dependency on other layers, thereby making those layers more competitive. NVIDIA, for example, partners with emerging cloud providers like CoreWeave and Lambda

Labs, creating competition for Azure, AWS, and Google Cloud. Meanwhile, cloud providers are developing their own chips, and model firms like xAI are building data centers to reduce reliance on large cloud vendors. These cross-layer moves reflect strategic hedging that can increase contestability across the value chain.

Finally, cross-stack investment and expansion can trigger competitive responses from incumbents, reinforcing their incentives to innovate. For instance, Google's research in generative AI was largely theoretical until the release of ChatGPT by OpenAI. In response, Google accelerated the commercialization of its own offerings, launching consumer-facing products such as Duet AI, Gemini, and AI-enhanced search. Ironically, many of the core machine learning breakthroughs that enabled this shift, such as the transformer architecture, had been developed by Google researchers themselves.

A lot of the weight on the scale of positive and negative effects of vertical integration will be taken up by user adoption and demand characteristics. We have touched above on how the economic fundamentals of GenAI models (and the applications currently built on top) do not seem to replicate the red alerts of the past. But another important question is: how are end users responding to this rapidly evolving landscape? To address this, we turn to a new survey of organizations using GenAI technologies. The survey helps illuminate how adoption is playing out across firms, the extent of multi-homing, and the perceived barriers to greater use, thus providing a demand-side view that complements the structural trends described above.

3 Adoption Dynamics

To understand how vertical integration and concentration at the infrastructure and model layers affect users, we conducted a survey of three hundred and twenty-three qualified respondents across a range of industries and geographies. The survey provides a demand-side perspective on the GenAI stack, offering insights into how organizations are adopting these tools, the relative prominence of different suppliers, and what barriers to broader deployment remain.

Respondents were recruited in May 2025 through the mailing list of the Digital Data Design Institute at Harvard (D^3), leading to a 0.5% response rate.[25] The majority of respondents report high seniority within their organization (32% are members of board, C-

[25]The study was reviewed and deemed exempt from IRB oversight by Harvard's Institutional Review Board. The survey tools are available upon request. The survey was sent to potential participants (individuals who signed up to the mailing list of the Digital Data Design Institute at Harvard) on April 29, 2025. We exclude respondents who were deemed ineligible and those who did not complete the survey in its entirety. The data described in this paper are all valid answers received by June 3, 2025.

suite, presidents, or executive officers; 15% are general managers or directors; and another 12% are senior managers). 58% are based in the US, followed by Europe (18%) and Asia (11%). Their organizations span multiple industries, from education (20%) and technology (20%) to professional services (17%), finance (10%), and health (8%).

As Table 1 shows, adoption of GenAI is widespread: nearly 90% of respondents report some use of GenAI tools in their organizations. However, usage is still considered to be too limited, constrained by a set of recurring challenges. The most frequently cited barriers to GenAI adoption include limited internal expertise, difficulty identifying clear use cases, concerns around data privacy and security, and cultural resistance to change. Many organizations noted that, while experimentation was common, sustained integration into core workflows remained a work in progress.

One of the most striking findings is the extent of multihoming. Most respondents reported using multiple GenAI models or tools simultaneously, especially combinations involving ChatGPT, Microsoft Copilot, Claude, and Gemini. Fewer than 20% of respondents said their organization relied on a single GenAI system, although, among those that did single home, more than half reported using ChatGPT.

Table 2 shows that multihoming is common across most industries represented in the survey. Technology and finance services report the highest levels of multihoming, with a large majority of respondents in these sectors indicating the use of multiple GenAI tools. Healthcare and manufacturing also show notable levels of multihoming, though slightly lower. This consistent pattern across all industries with sufficient responses suggests that the use of multiple GenAI systems is a widespread practice, not confined to any one sector.

A similar pattern appears in cloud infrastructure, though the prevalence of multihoming is lower: almost half of firms use more than one provider, likely to balance performance, price, or data locality needs. Regardless of multi-homing, the major providers remain Azure, AWS, and Google Cloud.

This behavior suggests a growing user awareness of vendor lock-in risks and the volatility of GenAI tools. It also underscores a key point from the supply-side analysis: despite emerging concentration at the top of the stack, users are actively managing that risk by diversifying their dependencies. The resulting picture is one of experimentation, cautious deployment, and strategic hedging, rather than full commitment to any single solution or provider.

Table 1: Survey Statistics

	Percent	Respondents
Usage		323
GenAI usage	87.0%	
Not enough GenAI usage	57.3%	
Barriers to adoption		193
Limited internal expertise	59.6%	
Unclear use cases	44.0%	
Data privacy and security	40.4%	
Resistance to change	37.8%	
Use cases		281
Summary of information	79.7%	
Writing assistance	75.1%	
Data analysis and querying	66.9%	
Short-form written content such as emails	65.5%	
Foundation models		281
Singlehoming	19.6%	
Multihoming	80.4%	
2 models	19.9%	
3 models	22.1%	
4+ models	38.4%	
Cloud infrastructure		171
Singlehoming	52.6%	
Multihoming	47.4%	
2 cloud providers	30.4%	
3 cloud providers	9.9%	
4+ cloud providers	7.0%	

Source: Authors' analysis based on an original survey on GenAI usage. Multihoming is defined as the use of GenAI solutions powered by different underlying foundation models. In the context of cloud infrastructure, multihoming refers to the use of services from multiple distinct cloud providers.

4 Conclusions

The evolution of GenAI is as much a strategic story as it is regulatory and technological. For business leaders navigating this landscape, several clear themes emerge from our analysis.

First, the GenAI market remains contested, especially at the foundation models and application layers. Unlike Web 2.0, the economic mechanisms that led to early market tipping (strong network effects, data feedback loops, and zero pricing) are currently weaker or absent. This presents a window of opportunity for firms to enter, differentiate, and build new value propositions. How long the window will remain open, and whether vertical integration may

Table 2: Multihoming by Industry

Industry	**Multihoming Share**
Technology	91.2%
Finance	87.5%
Government	83.3%
Education	78.9%
Professional Services	77.4%
Health	76.2%
Manufacturing	68.8%
Average across All Respondents	**80.4%**

Source: Authors' analysis based on an original survey on GenAI usage. Multihoming is defined as the use of GenAI solutions powered by different underlying foundation models. Only industries with a sufficient number of respondents are displayed in the table.

narrow the opening, remains to be seen. But, for now, there is little indication of a tendency towards monopolization or tipping.

Second, vertical integration is reshaping the AI supply chain. Both large established firms (GAMMAN) and other players are seeking integrations and partnerships that can extend their reach and sustain their growth as the industry develops. Firms that can integrate or partner across layers (such as compute, model development, and applications) will likely be better positioned to secure strategic advantages. In addition, strategic partnerships and mergers can have positive effects on innovation by allowing innovators to internalize knowledge spillovers, smoothen procurement and supply relationships, and stimulate competitive responses by the rest of the industry. Managers should assess where in the stack their organization can gain leverage and where partnerships might mitigate dependency and reduce costs.

Third, our survey results show that most organizations are experimenting with multiple GenAI tools and providers. This multihoming behavior reflects not just caution, but a strategic response to a fast-moving and uncertain market. Business leaders should view this as an opportunity to test, learn, and hedge, while avoiding early lock-in to any single vendor.

Business leaders will want to pay attention to the evolution of the industry in terms of trends in the cost of training models, how AI is being bundled into existing products, as well as the growing role of regulatory compliance which may increase as well as decrease entry and growth in GenAI. The GenAI market is not yet closed. It is contested ground, and there is still room to shape its future.

References

Miguel Antón, Florian Ederer, Mireia Giné, and Martin Schmalz. Innovation: the bright side of common ownership? *Management Science*, 71(5):3713–3733, 2025.

Dirk Bergemann and Alessandro Bonatti. Markets for information: An introduction. *Annual Review of Economics*, 11(1):85–107, 2019.

Andrei Hagiu and Julian Wright. Data-enabled learning, network effects, and competitive advantage. *The RAND Journal of Economics*, 54(4):638–667, 2023.

Andrei Hagiu and Julian Wright. Artificial intelligence and competition policy. *International Journal of Industrial Organization*, 2025. Forthcoming.

Michael L Katz and Carl Shapiro. Network externalities, competition, and compatibility. *The American economic review*, 75(3):424–440, 1985.

Tim O'Reilly. What is web 2.0: Design patterns and business models for the next generation of software. *Communication & Strategy*, (1), 2007.

Geoffrey G Parker and Marshall W Van Alstyne. Two-sided network effects: A theory of information product design. *Management science*, 51(10):1494–1504, 2005.

Jean-Charles Rochet and Jean Tirole. Two-sided markets: a progress report. *The RAND journal of economics*, 37(3):645–667, 2006.

Ilia Shumailov, Zakhar Shumaylov, Yiren Zhao, Yarin Gal, Nicolas Papernot, and Ross Anderson. The curse of recursion: Training on generated data makes models forget. *arXiv preprint arXiv:2305.17493*, 2023.

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. *Advances in neural information processing systems*, 30, 2017.