Regulating Competition in Generative AI: A Matter of Trajectory, Timing and Tools

Friso Bostoen[*] and Anouk van der Veer[†]

As antitrust scholars working on emerging technologies, we welcome the opportunity to contribute to the consultation on competition and generative AI (GenAI) organized by the European Commission (EC). In our submission, we tackle the question whether GenAI is a platform technology that requires immediate regulatory scrutiny to stay competitive. To do so, we address three sub-questions. The first concerns the trajectory of GenAI: is the market destined to evolve into a few ecosystems ruled by dominant platforms? (Section I) The second concerns the timing of regulation: given the new and quickly evolving nature of GenAI, what is the 'right' time to act? (Section II) The third concerns the regulatory tools: can we put existing competition law to work, or do the market dynamics call for a tailored instrument? (Section III)

The sections of our submission correspond to questions put forward in the consultation as indicated below (with questions that we address only in passing left out).

I. Trajectory
1. What are the main **components** (i.e., inputs) necessary to build, train, deploy and distribute GenAI systems?
2. What are the main **barriers to entry** and expansion for the provision, distribution or integration of GenAI systems and/or components, including AI models?
3. What are the main **drivers of competition** for the provision, distribution or integration of GenAI systems and/or components, including AI models?
9. Do the **vertically integrated companies**, which provide several components along the value chain of GenAI systems, enjoy an advantage compared to other companies?

III. Tools
4. Which **competition issues** will likely emerge for the provision, distribution or integration of GenAI systems and/or components, including AI models?
11. Do you expect the emergence of GenAI systems and/or components, including AI models to trigger the need to adapt EU legal **antitrust concepts**?
12. Do you expect the emergence of GenAI systems to trigger the need to adapt EU antitrust **investigation tools and practices**?

II. Timing adds important context to these questions/their answers.

[*] Assistant Professor of Competition Law & Digital Regulation, Tilburg University. Correspondence: f.bostoen@tilburguniversity.edu. Note that a version of this submission is to be published in Concurrences.
[†] Ph.D. Researcher, European University Institute. Correspondence: anouk.vanderveer@eui.eu.

I. The Trajectory of Generative AI

ChatGPT is synonymous with GenAI in the same way that Google has become a synonym for web search. Having acquired 100M users in just two months, it is considered the fastest-growing consumer product of all time (for comparison, it took Facebook four and a half years to acquire as many users).[3] But, in an evolution that is characteristic of digital markets, GenAI did not stop at being a product—instead, it evolved into a platform. The models have been growing larger and larger, but that trend may be starting to reverse. In this section, we look at both developments: the platformization of GenAI, and the trends in model size.

a. From Products to Platforms

It is helpful to think about the GenAI industry in layers. ChatGPT is just the tip of the iceberg. Such consumer-facing applications rely on GenAI models, in particular Large Language Models (LLMs)/Foundation Models (FMs). For example, ChatGPT—at least the paid version—is built on OpenAI's latest model, GPT–4. These models need to be trained, which requires two key inputs: data and compute.[4] OpenAI's GPT models were trained using compute provided by Microsoft's Azure cloud infrastructure, which is, in turn, powered by Nvidia's microchips (in particular, its graphics processing units or GPUs).[5] The training data was collected from the open internet, licensed from third parties, and provided by users.[6]

One can thus distinguish three GenAI layers: infrastructure, in particular compute; models; and applications, which is how models are 'deployed' to end-users. The model layer, in particular, shows platform characteristics, as independent developers can build their own products on top of this technological foundation. To facilitate this, OpenAI puts a range of application programming interfaces (APIs) at the disposal of developers, which can 'call' on the APIs in their own products.[7] A GPT Store has been set up for developers to sell their apps to users.[8] Meta goes a step further: it open-sourced its LLM, Llama, so that developers can customize the model itself.[9]

The GenAI market has structured itself not just around platforms but into ecosystems. 'Ecosystem' is a strategic management concept that refers to 'a group of interacting firms that depend on each

---

[3] Jon Porter, 'ChatGPT continues to be one of the fastest-growing services ever' (*The Verge*, 6 November 2023) <https://www.theverge.com/2023/11/6/23948386/>.

[4] A third key input is talented personnel.

[5] John Roach, 'How Microsoft's bet on Azure Unlocked an AI Revolution' (*Microsoft News*, 13 March 2023) <https://news.microsoft.com/source/features/ai/how-microsofts-bet-on-azure-unlocked-an-ai-revolution/>. More generally, see 'Computational Power and AI' (*AI Now Institute*, 27 September 2023) <https://ainowinstitute.org/publication/policy/compute-and-ai>.

[6] 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>.

[7] See < https://platform.openai.com/>.

[8] See <https://chat.openai.com/gpts>.

[9] See <https://llama.meta.com/>. Nuancing the open-source nature of Llama, see Charlie Hull, 'Is Llama 2 open source? No – and perhaps we need a new definition of open…' (*OpenSource Connections*, 19 July 2023) <https://opensourceconnections.com/blog/2023/07/19/is-llama-2-open-source-no-and-perhaps-we-need-a-new-definition-of-open/>.

other's activities', usually for the purpose of innovation.[10] The terminology puts the focus on the complex relations between players, as well as the environment in which interactions take place. Ecosystems tend to be orchestrated by a core platform, such as those found in the model layer. But what is notable in GenAI ecosystems is the outsized role played by infrastructure providers.

The outsized role of providers of infrastructure, in particular cloud computing, is explained in part by its concentrated nature. Microsoft Azure and Amazon Web Services (AWS) are the market leaders, with a combined 60–80% share of the market, with Google Cloud Platform (GCP) a distant third.[11] The chip industry is potentially even more concentrated: Nvidia is said to capture 90% of the AI GPU market.[12] Investors report that their portfolio companies spend up to 80% of their capital on compute resources, which underlines the importance of cloud infrastructure.[13] This is aggravated by the fact that compute is currently constrained, mainly due to the underlying need for chips.[14] For these reasons, cloud providers play a structuring role in GenAI ecosystems.

[10] Michael Jacobides, Carmelo Cennamo and Annabelle Gawer, 'Towards a Theory of Ecosystems' (2018) 39 Strategic Management Journal 2255.

[11] The exact market shares vary per country (and in reality, there is unlikely to be a single 'cloud' market in antitrust terms), see Authority for Consumers & Markets, 'Cloud Services' (Market Study, 2022) ACM/21/050317, 33–35; Autorité de la concurrence, Opinion 23-A-08 on Competition in the Cloud Sector, 29 June 2023, paras 296–304; Ofcom, Cloud Services Market Study (Final Report, 2023), para 11.30 and Annex 1.

[12] 'NVIDIA has 90% of the AI GPU Market Share' (*techovedas*, 21 January 2024) <https://techovedas.com/nvidia-has-90-of-the-ai-gpu-market-share-1-5-to-2-million-ai-gpus-to-be-sold-by-nvidia-in-2024/>. Note that Google has developed its own Tensor Processing Units (TPUs), which are optimized for AI training and can compete with GPUs, see <https://cloud.google.com/tpu>. Another significant breakthrough is Groq's Language Processing Unit (LPU), see <https://wow.groq.com/>.

[13] Guido Appenzeller, Matt Bornstein and Martin Casado, 'Navigating the High Cost of AI Compute' (*Andreessen Horowitz*, 27 April 2023) <https://a16z.com/navigating-the-high-cost-of-ai-compute/>.

[14] The main cloud players therefore invest strategically to safeguard access, as when Microsoft invested in the cloud infrastructure startup CoreWeave, which has preferential access to chips from NVIDIA (another of its investors), see Mark Haranas, 'Microsoft's CoreWeave Deal "Adds AI Pressure" To AWS, Google' (*CRN*, 2 June 2023) <https://www.crn.com/news/cloud/microsoft-s-coreweave-deal-adds-ai-pressure-to-aws-google>.

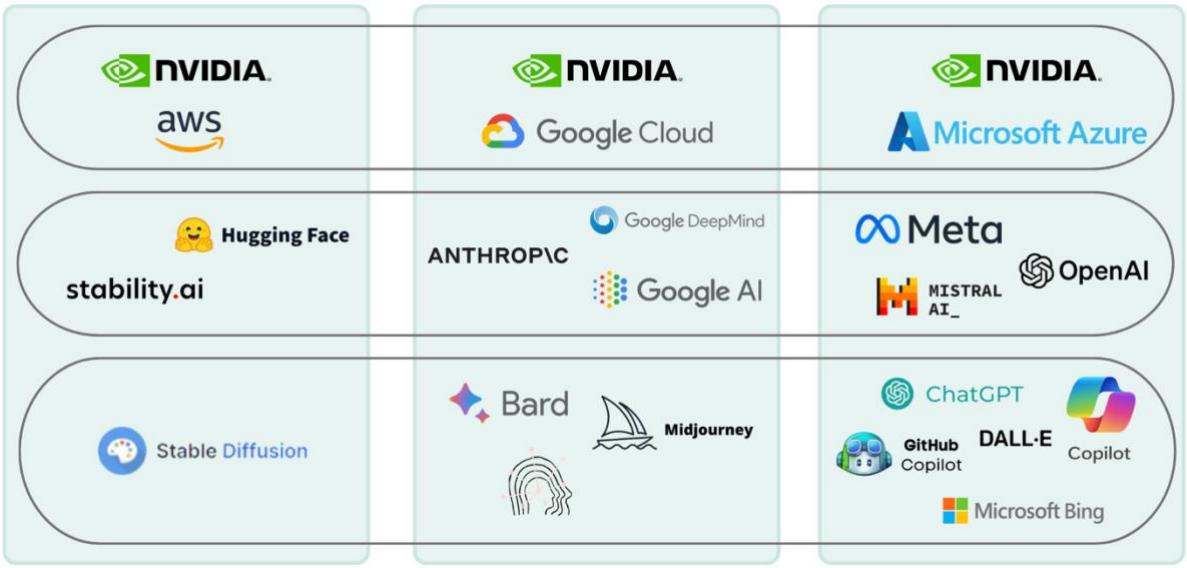

Figure 1. Non-exhaustive overview of GenAI ecosystems[15]

GenAI was hailed as a 'platform shift'.[16] The last platform shift occurred in the 2000s, which saw the rise of mobile (smartphones) along with the maturing of the web (search, social networking, e-commerce). This meant the rise of new firms like Google, Facebook (Meta) and Amazon, while older firms like Microsoft declined in relevance.[17] What is notable about the shift to GenAI is that the firms that rose to incumbency during previous shifts are also leading this one. This is due to their access to key inputs. As discussed above, Microsoft, Amazon and Google own the main cloud infrastructures. In addition, they have access to proprietary sets of training data; Google even has superior access to the open web.[18]

It is no surprise, then, that entrants in the GenAI space partner with infrastructure providers. OpenAI benefits from Microsoft's compute infrastructure via a $13B investment, for example.[19] Anthropic has received multi-billion investments from both Amazon and Google.[20] Of course, the infrastructure providers also have their own models (e.g., Google's Gemini/DeepMind). For now,

[15] The three layers (horizontal) group firms active at the levels of infrastructure, models and applications. The firms developing models—or platforms to collaborate on models, as with Hugging Face—are placed in the ecosystem (vertical) of their preferred cloud provider. The overview is based on public information (e.g., press releases), which may be incomplete.

[16] Ben Thompson, 'Windows and the AI Platform Shift' (*Stratechery*, 24 May 2023) <https://stratechery.com/2023/windows-and-the-ai-platform-shift/>.

[17] Apple, another company heralding the platform shift, was not exactly new but was certainly not (or no longer) an incumbent from the previous platform shift to desktop PCs.

[18] More websites allow its crawlers, which it has long used to build the web index for its search engine. Now, crawlers are used to collect training data, but websites can block them. Only 24% of news sites block Google's AI crawler, compared to 48% blocking OpenAI's crawler. See Richard Fletcher, 'How many news websites block AI crawlers?' (*Reuters Institute*, 22 February 2024) <https://reutersinstitute.politics.ox.ac.uk/how-many-news-websites-block-ai-crawlers>.

[19] Kyle Wiggers, 'Microsoft invests billions more dollars in OpenAI, extends partnership' (*TechCrunch*, 23 January 2023) <https://techcrunch.com/2023/01/23/microsoft-invests-billions-more-dollars-in-openai-extends-partnership/>.

[20] See <https://www.crunchbase.com/organization/anthropic/signals_and_news>.

there seems to be plenty of competition at the model level. Benchmarking is difficult, but every other week, the ranking of top models appears to change. At the time of writing, the French startup Mistral just announced that its latest model outperformed every other model except for GPT–4.[21]

b. Model Size: Two Roads Diverged into the Woods

Competition may seem fierce now, but where is the industry headed? So far, there has been a trend towards ever-larger models, as performance increases with size.[22] With about 1.8 trillion parameters, GPT–4 is over ten times larger than GPT–3.[23] Naturally, this trend raises concerns for competition authorities.[24] If there are large or even extreme returns to scale, as we saw with web 2.0 platforms, the market is bound for concentration. But authorities also noted that this is not an inevitability.[25]

In fact, there are signs of a reversal in model size. Smaller GenAI models are developed that rely on less trainable parameters but perform similarly or even better.[26] For example, DeepMind's Chinchilla[27] and Meta's Llama[28] have fewer parameters but outperform GPT–3.[29] Llama was, in turn, outperformed by the five-times-smaller Phi–2 model of Microsoft.[30] Stanford researchers managed to build Alpaca, a model on par with GPT–3.5, at a price of less than $600.[31] The

---

[21] Mistral AI Team, 'Au large' (press release, 26 February 2024) <https://mistral.ai/news/mistral-large/>.
[22] Jared Kaplan et al., 'Scaling Laws for Neural Language Models' (2020) arXiv:2001.08361.
[23] Maximilian Schreiner, 'GPT–4 architecture, datasets, costs and more leaked' (*The Decoder*, 11 July 2023) <https://the-decoder.com/gpt-4-architecture-datasets-costs-and-more-leaked/>
[24] The Portuguese competition authority focuses on the risks of the increasingly large models, see Autorida de da Concorrência, 'Competition and Generative Artificial Intelligence' (Issues Paper, 2023). Other authorities are still examining the market, see Autorité de la concurrence, 'Generative artificial intelligence: the Autorité starts inquiries ex officio and launches a public consultation' (press release, 8 February 2024) <https://www.autoritedelaconcurrence.fr/en/press-release/generative-artificial-intelligence-autorite-starts-inquiries-ex-officio-and-launches>; FTC, 'FTC Launches Inquiry into Generative AI Investments and Partnerships' (press release, 25 January 2024) <https://www.ftc.gov/news-events/news/press-releases/2024/01/ftc-launches-inquiry-generative-ai-investments-partnerships>; EC, 'Commission launches calls for contributions on competition in virtual worlds and generative AI' (press release, 9 January 2024) IP/24/85; Hungarian Competition Authority, 'GVH launches market analysis on the impact of artificial intelligence' (press release, 4 January 2024) <https://www.gvh.hu/en/press_room/press_releases/press-releases-2024/gvh-launches-market-analysis-on-the-impact-of-artificial-intelligence>.
[25] Competition & Markets Authority, 'AI Foundation Models' (Initial Report, 2023) paras 3.73–3.78.
[26] See, e.g., Yao Fu et al., 'Specializing Smaller Language Models Towards Multi-Step Reasoning' (2023) arXiv:2301.12726 and Ian McKenzie et al., 'Inverse Scaling: When Bigger Isn't Better' (2023) arXiv:2306.09479.
[27] Jordan Hoffmann et al., 'Training Compute-Optimal Large Language Models' (2022) arXiv:2203.15556.
[28] Hugo Touvron, 'LLaMA: Open and Efficient Foundation Language Models' (2023) arXiv:2302.13971.
[29] Massive Multi-task Language Understanding (MMLU) is a benchmark for AI models. Whereas GPT–3's accuracy is 53.9%, LLaMA's is 57.8% and Chinchilla's is 67.5%, see <https://paperswithcode.com/sota/multi-task-language-understanding-on-mmlu>.
[30] Phi-2 outperforms LLaMA on different benchmarks (though not all of them), see Mojan Javaheripi and Sébasatien Bubeck, 'Phi-2: The surprising power of small language models' (*Microsoft Research Blog*, 12 December 2023) <https://www.microsoft.com/en-us/research/blog/phi-2-the-surprising-power-of-small-language-models/>.
[31] Rohan Taori et al., 'Alpaca: A Strong, Replicable Instruction-Following Model' (2023) <https://crfm.stanford.edu/2023/03/13/alpaca.html>.

aforementioned Mistral model also offers an example: its compact 5GB size allows it to run on ordinary computers with about 16GB of RAM, while GPT–4 cannot run on consumer hardware.[32]

GenAI executives are seeing the writing on the wall. Microsoft's Gates predicts that GenAI models are reaching a performance plateau, so that GPT–5 will not be much better than GPT–4.[33] OpenAI's Altman believes models can still improve but believes these improvements will not come from increasing their size.[34] Such public statements should be taken with a grain of salt, as these parties have an interest in downplaying the importance of size so as not to attract regulatory scrutiny. However, an internal Google memo corroborates this story, setting out how the moat of large models is not as large as it may seem.[35]

In short, while there has been a trend towards ever-larger models, it is far from certain that this trend will continue. While it is early days, it is thus fair to say that the GenAI market does not *necessarily* exhibit the same tendency towards concentration as web 2.0 markets. Recognizing it is early in GenAI's trajectory brings us to the next issue: timing.

II. Regulatory Timing: Facing the Collingridge Dilemma

How does one regulate a technology—and, by extension, a market—that is in flux? That difficulty is illustrated by the Collingridge Dilemma.[36] Collingridge described how the consequences of a technology cannot be predicted early in its lifecycle—hence, the need for regulation is unclear. But by the time undesirable consequences manifest themselves, the technology has become part of the economic and social fabric, so that controlling it is extremely difficult—you can no longer put the genie back in the bottle. In other words, regulation is doomed to come either too early or too late.

In making the decision between regulating early or late, the spectre of the past looms large. In their calls for intervention, antitrust enforcers refer to the trajectory of the last platform shift. They describe how it started with revolutionary technologies disrupting markets, but ended with concentrated markets where firms solidified their dominance with anticompetitive practices. Antitrust intervention came too late—if it came at all. Federal Trade Commission (FTC) Chair Khan warns that 'public officials have a responsibility to ensure this hard-learned history doesn't

[32] See <https://docs.mistral.ai/models/>.
[33] Matthias Bastian, 'Bill Gates does not expect GPT–5 to be much better than GPT–4' (21 October 2023, *The Decoder*, 21 October 2023) <https://the-decoder.com/bill-gates-does-not-expect-gpt-5-to-be-much-better-than-gpt-4/>
[34] Will Knight, 'OpenAI's CEO Says the Age of Giant AI Models Is Already Over' (*Wired*, 17 April 2023) <https://www.wired.com/story/openai-ceo-sam-altman-the-age-of-giant-ai-models-is-already-over/>
[35] The memo is available at <https://www.semianalysis.com/p/google-we-have-no-moat-and-neither>. The memo includes statements like 'Giant models are slowing us down' and 'Focusing on maintaining some of the largest models on the planet actually puts us at a disadvantage'.
[36] David Collingridge, *The Social Control of Technology* (St. Martin's Press 1980).

repeat itself.'[37] EU Commissioner Vestager ties the timing of regulation even more closely to the trajectory of the technology:

> If the development of AI mirrors this history [of disruption followed by competition problems], then the right response is to phase in competition control gradually, in line with market growth. To respond faster than we did for Web 2.0, but to still give some time for the benefits of disruption to fully play out.[38]

But she is doubtful history will repeat itself. GenAI development is not originating from tinkerers' garages or university campuses as it did previously; instead, 'disruption from AI will come from within the nest of existing tech ecosystems.'[39] As pointed out above, this is due to the nature of GenAI models, which are trained using huge amounts of compute and data—a significant barrier to the entry of upstarts (certainly those without Big Tech partner). In short, the trajectory of GenAI may justify earlier intervention.

To dig a little deeper, let us consider the costs of early vs late intervention. As intervention in web 2.0 came (relatively) late, we have a better view of its downsides. The obvious one, as observed by Collingridge, is that remedying undesirable consequences becomes exceedingly difficult. We saw this clearly in the ineffectiveness of antitrust remedies.[40] The problem is explained in part by the particular nature of digital markets: due to economic characteristics such as network effects, digital markets tend towards concentration. This is not problematic in itself, as increasing scale benefits users. However, when firms then solidify their dominant position through anticompetitive conduct, it is difficult, if not impossible, to restore competition, as the remedy sagas of Microsoft and Google have shown.[41]

Early intervention also has its downsides. New technologies like GenAI have the potential to disrupt markets.[42] Regulation can arrest the process of dynamic competition. This can harm users who are denied the fruits of dynamic competition, whether they come in the form of a new product,

---

[37] Lina Khan, 'We Must Regulate A.I. Here's How.' (*The New York Times*, 3 May 2023) <https://www.nytimes.com/2023/05/03/opinion/ai-lina-khan-ftc-technology.html>.
[38] Margrethe Vestager, 'Making Artificial Intelligence Available to All—How to Avoid Big Tech's Monopoly on AI' (speech, 19 February 2024) SPEECH/24/931.
[39] ibid.
[40] See Friso Bostoen and David van Wamel, 'Antitrust Remedies: From Caution to Creativity' (2024) 14 Journal of European Competition Law & Practice 540.
[41] On Microsoft, see Daniel Gore and Ashwin van Rooijen, 'Ex Post Assessment of European Competition Policy: The Microsoft Cases' in Assimakis Komninos and Nicolas Petit (eds), *Ex Post Evaluation of Competition Cases* (Kluwer 2021) 17–44; *Google Shopping* has not benefitted from an objective retrospective, but complaints continue six years after the decision.
[42] For the classical work, see Clayton Christensen, *The Innovator's Dilemma: When New Technologies Cause Great Firms to Fail* (Harvard Business Review Press 1997). However, note, first, that GenAI is not necessarily starting at the low end of the market as foreseen by Christensen, due to the high cost of development and deployment compared to the technologies it is challenging (e.g., search). Second, it can be debated whether GenAI is a disruptive or rather a sustaining innovation.

a reshuffled market structure, or further innovation driven by that competition.[43] By contrast, regulation may benefit first movers who would like to see their early lead cemented. From that perspective, OpenAI's calls for regulation[44] can be seen as an effort to influence policymaking with the goal of solidifying its first-past-the-post position.[45] Asymmetrical regulation, which applies to market leaders but not challengers, can partially address the risk of setting in stone a particular market structure.

Importantly, GenAI models are not the only ones that learn from data—antitrust enforcers and policymakers do too. And that is perhaps the strongest argument against regulating (too) early: without deeper knowledge of its object, regulation is likely to miss the mark. But just like GenAI models, authorities are learning faster and faster. While it took a decade or longer before web 2.0 platforms faced scrutiny,[46] GenAI has been under investigation almost since day one. Sensibly, scrutiny tends to take the form of market investigations and other information-gathering exercises.[47] Having established a solid understanding of the industry's dynamics, authorities can then start taking a closer at potentially anticompetitive conduct.

At the same time, not every layer of the GenAI industry presents a mystery. Authorities started studying the cloud computing market years before GenAI entered the scene.[48] They did not stop there. The French competition authority followed up its market investigation with unannounced inspections at the offices of an unnamed graphics card manufacturer (read: Nvidia).[49] To ward off an EU probe, Microsoft changed its cloud licensing conditions, making it easier for customers to

[43] Microsoft, which quickly integrated GenAI features into its Bing search engine, did not win significant share from Google Search (less than 1 percentage point), see Jackie Davalos, 'Microsoft's Bing Market Share Barely Budged With ChatGPT Add-On' (*Bloomberg*, 18 January 2024). But the effects of innovation are more likely to show in the decreased use of search engines altogether.

[44] See, e.g., Cecilia Kang, 'OpenAI's Sam Altman Urges A.I. Regulation in Senate Hearing' (*The New York Times*, 16 May 2023) <https://www.nytimes.com/2023/05/16/technology/openai-altman-artificial-intelligence-regulation.html>.

[45] For example, requiring government licenses to train large-scale models, as suggested by Altman, could slow down progress to the benefit of the current market leader(s). More generally, any (symmetrical) regulation comes with compliance costs that larger firms can bear more easily. On the compliance costs of financial regulation, see Drew Dahl, Andrew Meyer and Michelle Neely, 'Scale Matters: Community Banks and Compliance Costs' (2016) The Regional Economist July 1.

[46] The EC's Google investigation marked the start of the platform antitrust, see EC, 'Commission probes allegations of antitrust violations by Google' (press release, 30 November 2010) IP/10/1624. Important reports followed some years later, see notably Autorité de la concurrence and Competition & Markets Authority, 'The Economics of Open and Closed Systems' (Report, 2014).

[47] See the consultations and investigations referenced above, footnotes 24–25.

[48] See the reports referenced above, footnote 11.

[49] Autorité de la concurrence, 'Le rapporteur général de l'Autorité de la concurrence indique qu'une opération de visite et saisie inopinée a été réalisée dans le secteur des cartes graphiques' (press release, 27 September 2023) <https://www.autoritedelaconcurrence.fr/fr/communiques-de-presse/le-rapporteur-general-de-lautorite-de-la-concurrence-indique-quune-operation>.

switch providers.[50] EU policymakers even felt comfortable enough to adopt the Data Act, which also regulates cloud providers, including by facilitating customer switching/multi-homing.[51]

As the interventions in cloud markets show, a similar issue can be tackled by different instruments. This throws up the question of regulatory tools: which type of intervention, if any, does the GenAI market call for?

III. Regulatory Tools: Of Hammers and Nails

'If the only tool you have is a hammer, everything looks like a nail.' This quote is sometimes (ab)used to criticize the EC for using antitrust enforcement to solve problems that are arguably not *competition* problems. Meanwhile, the EC's toolbox has expanded: it is now also the main enforcer of the DMA. This new tool certainly shows overlap with the abuse of dominance provision, but its enforcement mechanism is very different as it takes an *ex ante* rather than *ex post* approach.

Which tool is best for 'the job', i.e., safeguarding a competitive GenAI ecosystem? In answering this question, timing matters. Competition law is a flexible tool; its general provisions can immediately be applied to new technologies and market dynamics. By contrast, sectoral tools like the DMA must define their scope and conduct rules in greater detail. This is hard to do before competitive issues have clearly been identified, in particular via antitrust enforcement. Once in place, however, sectoral tools can be applied within a shorter timeframe than competition law, where investigations take years. Moreover, there may be issues of competition that escape competition law, which can justify additional regulation.

Let us start by surveying the applicability of competition law.

a. Competition law

For the GenAI industry, competition law's abuse of dominance and merger control provisions are most relevant.

Before finding any abuse, the authority must establish a dominant position. This appears difficult in the GenAI model layer, which boasts a variety of players. Even if we consider only large-scale models, there appears to be, at most, an oligopoly. The same goes for the infrastructure layer. In cloud computing, there are three large players without any one of them being clearly dominant. Based on the market shares quoted above, the firm most likely to qualify as dominant is Nvidia.

As of yet, the contours of abusive conduct are also unclear. Extrapolating from the current market structure and dynamics, however, the following types of abuse can be envisaged.

---

[50] Foo Yun Chee, 'Microsoft offers to change cloud practices to ward off EU antitrust probe' (*Reuters*, 28 March 2023) <https://www.reuters.com/technology/microsoft-offers-change-cloud-computing-practices-after-rivals-complaint-source-2023-03-28/>.

[51] Regulation (EU) 2023/2854 of the European Parliament and of the Council on harmonised rules on fair access to and use of data (Data Act).

- **Input foreclosure**. Compute is one of two key inputs of GenAI models. The main providers of compute also have their own models. Therefore, they may have an *incentive* to foreclose the developers of competing models from this input. But their *ability* to foreclose may be limited, since there are at least two other major infrastructure providers. Nevertheless, model developers may want to pre-empt such a move by partnering with an infrastructure provider, which they have been doing widely.
- **Exclusivity**. The other key input for GenAI models is data. Given the extensive datasets that GenAI models are trained on, however, no one firm holds a significant enough amount of data to foreclose. Indeed, models are trained on the open web, or at least a chunk of it (e.g., the open CommonCrawl index). A competitive advantage may be gained by adding valuable, proprietary datasets to the training data. Hence, a model developer with deep pockets may try to lock up those datasets through exclusivity agreements.[52]
- **Leveraging**. The platformization of GenAI models opens the door to leveraging, in particular from the platform to the application layer. A model provider could, for example, favor its own apps over those of independent developers that rely on its model. This could take a variety of forms, from demoting in ranking to degrading API access.
- **'Open early, closed late'**. Open-source models play an important role in the GenAI landscape by allowing developers with fewer means to nevertheless build on powerful models. Having reaped the rewards of openness, a provider could start closing down its model. Such an 'open early, closed late' strategy need not concern models; it could also relate to other inputs in the GenAI chain.[53]

Merger control is designed to prevent the creation of a dominant position via acquisition. It thus plays an important proactive role. The GenAI model market has not seen large acquisitions.[54] It has seen a wave of partnerships, the most eye-catching being Microsoft–OpenAI. But the EU Merger Regulation only applies when there is a change of 'control', meaning one firm acquires decisive influence over the other.[55] The Microsoft–OpenAI partnership appears structured to avoid such influence (Microsoft has a minority share of 49%), though the leadership struggle at OpenAI

[52] Data from Reddit, for example, is said to be particularly useful for model training. Google inked a deal with the company, though there is no mention of exclusivity, see Anna Tong, Echo Wang and Martin Coulter, 'Reddit in AI content licensing deal with Google' (*Reuters*, 22 February 2024) <https://www.reuters.com/technology/reddit-ai-content-licensing-deal-with-google-sources-say-2024-02-22/>.
[53] For example, Microsoft has threatened to cut off access to its web index to firms that use it to compete with its own AI chatbots, see Leah Nylen and Dina Bass, 'Microsoft Threatens Data Restrictions in Rival AI Search' (*Bloomberg*, 25 March 2023) <https://www.bloomberg.com/news/articles/2023-03-25/microsoft-threatens-to-restrict-bing-data-from-rival-ai-search-tools>.
[54] The infrastructure layer has seen such acquisitions, for instance of Arm by Nvidia. The transaction was called off after regulatory scrutiny and opposition, see Federal Trade Commission, 'Statement Regarding Termination of Nvidia Corp.'s Attempted Acquisition of Arm Ltd.' (press release, 14 February 2022) <https://www.ftc.gov/news-events/news/press-releases/2022/02/statement-regarding-termination-nvidia-corps-attempted-acquisition-arm-ltd>.
[55] Council Regulation (EC) No 139/2004 of 20 January 2004 on the control of concentrations between undertakings (EC Merger Regulation), article 3.

demonstrated Microsoft's significant behind-the-scenes influence.[56] As a result, several competition authorities are probing the partnership.[57] Their vigilance is certainly welcome, as there is a risk that the disruption brought by GenAI is 'co-opted' by incumbents.[58]

b. Digital Markets Act

GenAI models are not on the DMA's list of 'core platform services' (CPSs) and are therefore not subject to its rules.[59] Other layers of the GenAI chain are, to varying degrees, subject to the DMA. GenAI applications, such as chatbots, may be subject to the DMA as 'online search engine' or 'virtual assistant'. As for the infrastructure layer, 'cloud computing services' are on the list CPSs. But at the time of writing, the EC had not designated any cloud providers as gatekeepers. Their omission was notable,[60] and is likely due to definitional and user number difficulties.[61] Cloud providers may still be designated following a market investigation, which some countries are reportedly pushing for.[62] At the same time, they are already regulated by the aforementioned Data Act.

The European Parliament has called for a market investigation to examine whether GenAI services need to be added to the list of CPSs.[63] Is the regulation of GenAI under the DMA, beyond its current coverage, justified? Conceptually, if the GenAI model market turns out to be a natural monopoly (with one large provider) or evolves into an oligopoly (with, say, 2–4 providers), there may be a case for sectoral regulation, as antitrust law is ill-suited to either situation.[64] But the current market dynamics do not clearly point towards that outcome.

This relates to the timing question discussed above. New markets are not an easy or even suitable target for sectoral regulation. This is in part because such regulation demands a deep understanding

---

[56] See Charles Duhigg, 'The Inside Story of Microsoft's Partnership with OpenAI' (*The New Yorker*, 1 December 2023) <https://www.newyorker.com/magazine/2023/12/11/the-inside-story-of-microsofts-partnership-with-openai>.
[57] Competition & Markets Authority, 'CMA seeks views on Microsoft's partnership with OpenAI' (press release, 8 December 2023) <https://www.gov.uk/government/news/cma-seeks-views-on-microsofts-partnership-with-openai> and Federal Trade Commission, 'FTC Launches Inquiry into Generative AI Investments and Partnerships' (press release, 25 January 2024) <https://www.ftc.gov/news-events/news/press-releases/2024/01/ftc-launches-inquiry-generative-ai-investments-partnerships>.
[58] See Mark Lemley and Matthew Wansley, 'Coopting Disruption' (2024) <https://ssrn.com/abstract=4713845>.
[59] Regulation (EU) 2022/1925 of the European Parliament and of the Council on contestable and fair markets in the digital sector (Digital Markets Act), articles 2(2) and 3(1).
[60] The only other CPS category without designation is virtual assistants, see EC, 'Digital Markets Act: Commission designates six gatekeepers' (press release, 6 September 2023) IP/23/4328.
[61] On those difficulties, see Friso Bostoen 'Understanding the Digital Markets Act' (2023) 68 Antitrust Bulletin 263, 275–277.
[62] Luca Bertuzzi, 'Are EU regulators ready for concentration in the AI market?' (*Euractiv*, 3 November 2023) <https://www.euractiv.com/section/artificial-intelligence/news/are-eu-regulators-ready-for-concentration-in-the-ai-market/>.
[63] European Parliament, Resolution of 16 January 2024 on competition policy – annual report 2023 (2023/2077(INI)).
[64] Oversimplifying: In case of natural monopoly, the monopoly is efficient and should be regulated as such—not through the introduction of competition. In case of oligopoly, no single firm may be dominant, and they may not even be 'collectively dominant', but their conduct may nevertheless be anticompetitive.

of the market dynamics and competition issues, which requires (enforcement) experience and, therefore, time to acquire. In that sense, Scott Morton pointed out that the DMA's rules 'only apply to fairly old technology, like search and operating systems and social networks, that have been around for 15 years or more, so we understand them reasonably well.' The DMA itself has a built-in time lag: before being designated a gatekeeper, a firm must enjoy an 'entrenched and durable position', which means its CPS must have had 45 million EU users for *at least three years*.[65]

IV. Conclusion

We sought to contribute to the ongoing debate on competition in GenAI by, in the first place, asking the right questions. When it comes to regulation, we believe those questions relate to the trajectory of GenAI, the timing of intervention, and the regulatory tools relied on. Answering these questions is more difficult. First, GenAI has followed a clear platformization trajectory. It is unclear, however, whether that platformization will also end in significant concentration. The trend towards ever-larger models points in that direction, but there are signs of a reversal. Second, in terms of timing, the choice is essentially between regulating too early and too late. As enforcers have erred towards the latter in web 2.0 markets, they appear determined to get in early now. But intervention requires understanding, which is still limited. One strategy to solve this dilemma is to progressively dial up enforcement, starting with information-gathering initiatives. Third, the available tools range from competition law to more sectoral regulation like the DMA. Each has their benefits (and costs), and the choice between them ties in with timing. The DMA itself only came after widespread antitrust scrutiny, and we do not recommend skipping that step now.

[65] Digital Markets Act, article 3(2)(c).