

Kluwer Competition Law Blog

Decoding Competition Concerns in Generative AI

Kalpana Tyagi (University of Maastricht) · Friday, June 27th, 2025

As per [McKinsey](#), a top consulting firm, generative AI could add \$2.6 trillion to \$4.4 trillion to the global economy within the next decade. This emerging value and the growth potential of generative AI has also attracted the attention of competition authorities worldwide. In light of their experience in digital markets, competition authorities considered that an early intervention in the market may help avert competition concerns, such as irreversible tipping of the markets.

Key considerations and concerns revolve around high barriers to market entry, network effects and tipping, integration of foundation models into the larger ecosystem of digital gatekeepers and dependency on large digital market players for essential infrastructure to train generative AI models. The [UK Competition and Markets Authority \(CMA\)](#) studied competition concerns along the foundation models (FM) value chain, and acknowledged competition concerns such as self-preferencing, vertical integration, refusal to supply and unfair terms and conditions. The [French Competition Authority \(FCA\), Autorité de la concurrence](#), assessed competition concerns at the following three layers in the generative AI value chain: the upstream layer, as in the AI infrastructure layer; the middle layer, i.e. the AI modelling layer and the AI deployment layer. This discussion follows this three-layered approach to assess competition concerns in the generative AI value chain and offers an overview of the key insights and research findings of the pre-print “Mapping competition concerns along the generative AI value chain” available [here](#).

The upstream infrastructure layer

In the upstream market, the key inputs are computing power, data and skilled workforce. In addition, availability of venture capital and financial resources is also essential to a vibrant generative AI market.

Considering the large sunk costs involved in developing generative AI models, a vibrant and well-functioning capital market is critical to a local innovative ecosystem. In fact, a key reason for the thriving generative AI market in Silicon Valley is the abundant availability of venture capital (VC) to fund promising start-ups.

Flow of VC funding is also an indicator of the potential and promise of an emerging market. As Google became the dominant search engine, investment in search engines virtually came to a halt. The market for funding R&D and commercialization in generative AI, on the other hand, saw a geometric increase. Between 2022 and 2024, funding in generative AI-based start-ups quadrupled

to reach [\\$25.2 billion in 2024](#). While it is true that financial funding is the lubricant that helps run smoothly high sunk cost-driven innovative projects, considering that this limitation in the upstream layer cannot be remedied through competition and regulation, the follow-on discussion revolves around the three other key inputs that are *sine qua non* for generative AI. These include: computing power, data and talent acquisition (or acqui-hires).

Computing power

Central Processing Units (CPUs) are the brain of our digital devices, such as laptops and computers. A key limitation with CPUs is that they cannot cope well with the complexity of large language, deep learning models. To resolve this, graphic cards, used mainly for gaming, emerged as a promising alternative. Graphic processing units (GPU) could work well with complex, high-powered processing required in training large language models. Thus, the rise of big data revolution followed by improvements in hardware capabilities, including innovations in GPUs, led scientists to pro-actively engage in neural network-based approach to training generative AI models.

Nvidia was an important contributor to this new wave of innovation. As ChatGPT, one of the fastest and most widely adopted app in digital history, trained on Nvidia's offerings, the latter's valuation quickly rose from \$300 billion in November 2022 to [\\$2.3 trillion in 2024](#). From an innovation perspective, this is an important consideration, as Nvidia is a remarkable contributor to the uptake of the generative AI revolution. However, it has also been under the [antitrust radar for its exclusionary practices](#) in the sale of GPU chips, as well as its acquisition of Run:ai, an AI optimization software.

Nvidia's capabilities are complemented by other key players along the manufacturing value chain. For example, it outsources manufacturing to [Taiwan Semiconductor Manufacturing Company \(TSMC\)](#) and [Samsung Electronics](#). TSMC is, in turn, dependent on the Dutch national champion, ASML, that is the world's 'only' supplier of extreme ultraviolet (EUV) lithography machines, that is an essential input to manufacture GPU chips. To develop an ecosystem around its core offering of GPU chips, Nvidia has developed a software, Computer Unified Device Architecture (CUDA) that helps developers simultaneously use processing power from different sources.

To further strengthen this software capability, [Nvidia also recently acquired Run:ai](#), a so-called killer acquisition. Nvidia/Run:ai merger was first caught by the Italian Competition Authority (ICA), L'Autorità Garante Della Concorrenza e Dell Mercato (AGCM) under its special call-in powers. As per [Article 16\(1-bis\) of the Italian competition law](#), three conditions must be cumulatively met for a proposed transaction to qualify for review: first, the notification to the merged entity must follow within the first six months of the consummation of the transaction; secondly, either the merger should meet the turnover-based thresholds under Article 16(1) or alternatively, the merged entity must have an annual worldwide turnover of €5 billion or more and thirdly, there must be an identifiable risk to innovation in part or entirety of the relevant market. As the merger met all the three conditions, the ICA referred the merger to the European Commission under Article 22(1) of the EU Merger Control Regulation. Following a phase I investigation, the Commission, [unconditionally approved the transaction](#).

Unlike Nvidia, however, the other players in the market, namely TSMC and ASML, remain

focussed on their core capabilities, and have not developed an ecosystem around their main offering. Nvidia, with over 90 per cent market share, is a monopolist in the market for GPU chips. However, it is not without competition. Key sources of potential competition include Google, Amazon and Microsoft, that are investing heavily in developing their own GPU chips. In addition, Meta is working on its eAccelerator chips and IBM is working on its flagship, Telum project to offer a viable competitive alternative to Nvidia's GPUs. Considering that multiple competition projects are being pursued to develop a better GPU, any one of these projects, if successful, may in the long run, prove to be a workable competitive alternative to Nvidia's GPUs.

While Nvidia's GPU may face some emerging competition, [the market for cloud computing is dominated by the three cloud hyperscalers](#), namely Amazon Web Services, Microsoft Azure and Google Cloud Platform. In light of the findings from its [cloud market study](#) on the position of strength enjoyed by these hyperscalers, the UK telecommunications regulator, OfCom, requested a full market investigation by the UK Competition and Markets Authority. As noted by the [French Competition Authority](#), and the [UK Competition and Markets Authority \(CMA\)](#), the market for cloud computing is highly concentrated on account of substantial costs involved in building a data center, technical barriers to customer migration and data integration, egress fees, committed customer spend discounts and restrictive licensing practices. The highly concentrated nature of digital markets and cross-linkages between different services in the ecosystem helps hyperscalers further fortify their position of strength. Increased demand for GPUs and cloud computing following the rise and the mainstream adoption of generative AI is only expected to strengthen the position of hyperscalers. This can be attributed to the fact that developing, running and maintaining generative AI models requires large computing resources.

Open AI, one of the notable unicorns in generative AI, benefited substantially from its close collaboration with Microsoft. In addition to the \$12 billion funding that OpenAI received from prominent Silicon Valley based investors, such as Microsoft, Khosla Venture, and A16Z, Microsoft also developed a dedicated cloud infrastructure for OpenAI to train its GPT models. Following this close collaboration in the infrastructure and modelling layer, OpenAI's products are exclusively available on Microsoft Azure. Thus, this collaboration has spill-over effects across the entire generative AI value chain. Considering the centrality of cloud infrastructure and the position of strength enjoyed by these hyperscalers, an important consideration is whether the Digital Markets Act (DMA) can offer relief. The consideration here is not whether generative AI can be captured by the DMA (an issue that I turn to later in the discussion), instead the question here is whether competition and contestability in cloud computing, that is 'an essential infrastructure' for generative AI, can be facilitated through the DMA. Trite to add here that Article 2(2)(i) of the DMA refers to cloud computing as a core platform service (CPS). The hyperscalers, Google, Amazon and Microsoft, have already been designated as gatekeepers under the DMA. Currently, however, their cloud computing services are not qualified as CPSs as they do not meet the prescribed thresholds under the DMA nor has the EC decided to trigger designation based on qualitative criteria.

Bearing in mind the centrality of cloud to the digital and now the generative AI infrastructure, it may be important to re-consider whether cloud services can be caught by and subject to the obligations under the DMA. The market investigation tools as available under Chapter IV of the DMA may be a useful starting point to evaluate this possibility. Instructive in this regard is also a competition law-related development, wherein Google complained to the European Commission that Microsoft was engaging in [anti-competitive tying](#) by forcing customers to select its cloud service, Azure alongside its main offering, Windows Server.

Data

Data is the food that feeds generative AI models. Data is thus, a key input in the training of generative AI models. Google, Amazon, Microsoft, Meta and Apple (GAMMA) control rich sources of valuable data. Google has access to the data generated by YouTube, Google Search, Gmail, Google Maps and its Android App Store. Meta has rich sources of personal communication and content generated data accessible on its social networking and communications apps, namely, Facebook, Instagram and WhatsApp. Amazon has access to users purchasing habits on Amazon.com, as well as data generated over and for its services, such as Prime, Twitch and Alexa. Microsoft has access to data generated on its productivity software, and communication apps, such as Teams. Apple has access to data on Apple Watch, Siri and its App Store.

In addition, considering their position of strength, large digital firms may also tend to benefit from the data feedback loop. The use of these models leads to more data input, which leads to a better product, and as this data is used to train the model, a better product leads to a larger user base, which in turn again offers more data, and the cycle continues perpetually.

Generative AI has . It was trained using ‘300 billion words’ scrapped from the worldwide web. This content included both personal as well as non-personal data, without the consent of the rightsholders or the data subjects. This draws a close parallel with Google’s advantage as the world’s most popular search engine, which, in turn, offered it an unparalleled advantage while training its generative AI model Gemini (formerly, Bard). Google derives unparallel advantages from being the world’s most popular search engine. It receives over 93 per cent of unique queries, when compared to 3.8% queries received by its closest competitor, Bing. In other words, thanks to its unique access to queries and data, Google’s algorithms can make more meaningful inferences about the relevance of a given query. Likewise, generative AI models, such as ChatGPT, have trained principally on web-scrapped data. Moreover, to restrict follow on access to data to other small and medium enterprises, ChatGPT has restricted access to this proprietary data generated on its platform.

Data can be of many different types. When original and a result of [author’s own intellectual creation](#), data source may be subject to copyright and related rights. From that lens, the outcome in the ongoing copyright related lawsuits in the US, EU and the UK, have important implications for competition and innovation in the market for generative AI models. In the EU, the [licensing framework](#) available under the 2019 Copyright in the Digital Single Market Directive (CDSM) is seen as an enabler to ensure lawful access to copyright-protected proprietary data for training generative AI models (a process known as [text and data mining](#)).

Thus, lawful access to copyright-protected data is not only about fair and proportionate remuneration of the romanticised human author that sits at the core of copyright; the issue is equally central to facilitating contestability in generative AI.

In case gatekeepers have a data advantage that flows from their CPS (such as Google’s data advantage from Google Search), access can be enabled via the DMA. Article 6(10) of the DMA requires authorized ‘effective high-quality, continuous and real-time access to, and use of, aggregated and non-aggregated data, including personal data’. Such access must be provided in compliance with the consent requirements as set in the 2016 General Data Protection Regulation

(GDPR). In addition, Article 6(11) requires access to ‘ranking, query, click and view data’ by the gatekeeper on ‘fair, reasonable and non-discriminatory’ terms. A stringent application of these provisions (on this also see, the [Google compliance reports](#)) can be an enabler of competition in training generative AI models. This may be an important enabler, as a dichotomy exists relating to training and control over data. While digital startups are data hungry, the rise of [large language models and synthetic data](#) has diminished Google’s (and other digital gatekeepers) data dependency. Google’s older algorithms required over a trillion examples to [make meaningful inferences, whereas its more recent algorithms](#) (powered by generative AI-driven advances such as word vectors and transformers) require only one billion examples to make a meaningful inference. Thus, while entrants in the generative AI sector remain data hungry due to lack of economies of scale and scope, incumbent digital players, such as Google and Microsoft (also consider their close partnership with OpenAI), not only sit as gatekeepers on large reservoirs of data, but there enhanced algorithms are also increasingly data efficient. Additionally, the design of smarter algorithms is facilitated by talented technical teams.

Talent acquisition (or acqui-hire)

Talented technical personnel are key to innovation in generative AI. In 2024, two Nobel Awards were awarded in [physics](#) and [chemistry](#) to researchers working in machine learning, artificial neural networks and the application of machine learning in drug discovery.

In 2021, over 65% of Machine Learning/AI researchers were hired by the industry. Such hiring though legally permissible, it can under circumstances fall foul of merger control rules, particularly when the acquirer [acqui-hires](#) the key personnel, such that it offers him *de facto* control over the acquired undertaking. The Microsoft/Inflection merger presented one such situation wherein Microsoft hired the key personnel from Inflection such that effective ownership was passed into the hands of Microsoft.

The UK CMA assessed the merger as it was deemed a [relevant merger situation](#) under the Enterprise Act, 2002. Following a phase-1 investigation, the CMA did not refer the transaction for further assessment, as the merger did not raise competition concerns in the market for the development and supply of foundation models. The European Commission initially accepted the referral from the EU Member States under Article 22 of the EU Merger Regulation. However, following the CJEU’s decision in [Illumina/Grail](#) that [Article 22 was ‘not a corrective mechanism to capture concentrations’ that escaped national merger control thresholds](#), the Member States withdrew their referral request and [the Commission chose not to proceed further with its assessment of the Microsoft/Inflection deal](#). Mergers such as Microsoft/Inflection and Nvidia/Run:ai highlight a ‘threshold’ gap in the current EU merger control framework as this is the first filter that determines the applicability of merger regulation.

The AI modelling layer & the AI deployment layer

Following the availability of key inputs, the next level is training in foundation models. Training an AI model is a complex and energy-intensive process. Quality of training and output therein

depends on the quality of available inputs at the infrastructure layer (as discussed above).

The trained models can be either general-purpose, or they can be vertically specialized foundation models targeted at certain niche segments in the industry. A GPAI, if available in open source, can be fine-tuned to perform particular tasks, using a process known as ‘domain or task specific fine tuning’. ChemBERTa, for example, is built on Meta’s LLama-2 models and is a vertical FM model targeted at chemistry, cheminformatics, and drug discovery. Considering the targeted nature of these vertical models, they perform better in their targeted niche compared to the GPAI.

Building a general-purpose AI model from scratch is an expensive process. As per estimates, developing a [GPAI from scratch can cost up to US\\$8.8 trillion](#). GAMMA’s FM models are more general purpose and can be trained for deployment across a range of verticals. These models can be open-source or close source. While a large majority of models, at least at the moment, are available open source, it may be important to consider that these firms do not follow the [open early, closed later strategy](#).

Following the foundation layers of AI infrastructure and AI training, comes the deployment layer. At the deployment layer, AI services are offered to the end user. Microsoft’s Co-Pilot, Meta’s AI-based answer engines on Facebook, Instagram and WhatsApp are some examples of the integration of foundation models into existing digital ecosystems to enhance user output.

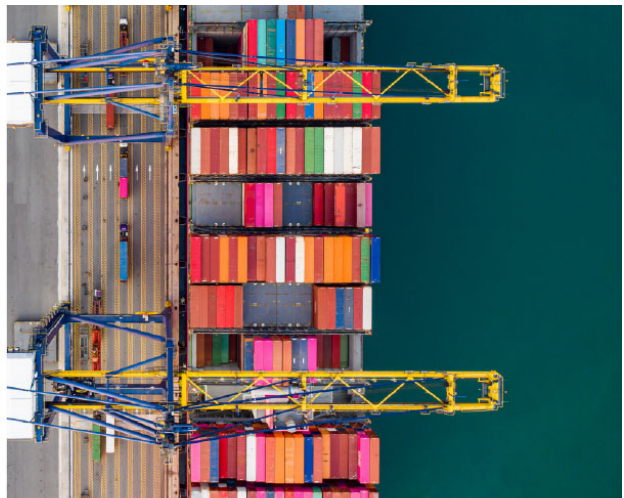
Timely application of the DMA can be a key enabler of competition at this level of the value chain. There are two possible approaches to capture generative AI under the DMA. A long-drawn approach may be the possibility to amend the DMA and offer an additional category of CPS looking only at the generative AI models. An alternative, and perhaps more efficient approach, as also illustratively emphasized by [Coeuré](#), [Ribera Martínez](#), [Picht](#) and [Hoppner](#), may be to capture generative AI as an embedded feature in the gatekeeper’s already-designated core platform services. A long drawn path to amend the DMA may unsettle the enforcement process, particularly in light of the increasing [geopolitical considerations in competition and regulatory enforcement](#). On the other hand, an effective and efficient use of the provisions of the existing regulatory framework, may not only align with recommendations of the [Draghi Report to streamline enforcement](#), but it may in addition, also facilitate a timely, proportionate and effective intervention in generative AI.

To make sure you do not miss out on regular updates from the *Kluwer Competition Law Blog*, please subscribe [here](#).

Now live

Global ESG Legal Compliance on VitalLaw®

Streamline compliance with ESG law and regulations with expert resources, global news updates, and legal research tools.



[Learn More →](#)

This entry was posted on Friday, June 27th, 2025 at 10:00 am and is filed under [AI](#), [Source: OECD](#)“>[Competition](#), [Data protection](#), [Regulation](#)

You can follow any responses to this entry through the [Comments \(RSS\)](#) feed. You can leave a response, or [trackback](#) from your own site.