oxera

AI, competition, and regulation: navigating the challenges ahead

—

Briefing paper for the OEC meeting of 11 June 2025

04 June 2025

Contents

Figures and tables

1 Introduction

Artificial intelligence (AI) is a general-purpose technology, comparable in its transformative potential to innovations such as the steam engine in the 18th century or the personal computer in the mid-20th century. It is driving technological change at a rapid pace, reshaping industries, markets, and broader economic structures. While AI offers the prospect of significant productivity gains and economic growth, these benefits are not guaranteed and come with inherent risks and costs. Among the risks are the potential for amplifying inequality in society, increasing the spread of misinformation, increasing cybercrime, copyright infringement risks, disruption of the labour market, and erosion of privacy. Consequently, policymakers around the world are actively debating how to balance the benefits that AI offers against the potential risks, considering various interventions.

This briefing paper focuses on a specific part of AI's impact on society: the dynamics of competition within 'AI markets', i.e. the markets directly involved in producing and deploying AI (e.g. AI chips, datacentres and model training) and the adjacent markets within which AI is being used (i.e. services and products that use AI models).[1,2] Our analysis pays particular attention to generative AI, and its interaction with the established market positions of major digital technology firms. The increasing scrutiny of AI markets by competition authorities—evidenced by published reviews and ongoing legal cases—raises the question of whether this regulatory activity represents an adequate response to concrete risks likely to impact competition, or, on the contrary, is a potential overreaction prompted by past criticisms of delayed intervention in other digital markets.

Against this backdrop, the Oxera Economics Council (OEC) meeting on 11 June 2025 will address the following key questions.[3]

1. Is there a demonstrable need for intervention in AI markets to safeguard competition and, if so, at what level(s) of the value chain?
2. Do competition authorities have the power to regulate competition in AI markets effectively, or are new tools and approaches required?
3. If intervention is deemed necessary, should this be done 'early on'?

In this briefing paper we deliberately do not take a position on these issues, and instead focus on setting out a list of key topics to facilitate an informed discussion with members of the OEC. The questions that we are proposing to discuss in detail are presented at the end of Sections 3, 4 and 5—and collected in Section 6.

The remainder of this briefing paper is organised as follows: Section 2 provides essential background information on AI markets; Section 3 examines the key economic features of these markets and the principal theories of harm currently being considered in relation to incumbent

[1] Following the recent OEC meeting on 11 June 2025, this briefing paper has undergone minor revisions to enhance its clarity and address any typographical errors.
[2] While some broader economic and societal impacts of AI are occasionally mentioned in this paper—as they interrelate with competition issues (e.g. misinformation, intellectual property, erosion of privacy, lack of governance, labour market disruption, macroeconomic effects, etc.)—they are outside the scope of this briefing paper.
[3] Some of these questions are topics on which Oxera is currently working on, on the behalf of its clients.

tech firms; Section 4 considers the adequacy of existing powers held by competition authorities to regulate AI markets (in case that intervention should be deemed necessary); Section 5 examines the question of the right time to intervene; and Section 6 sets out the questions for discussion for the OEC members.

2 Background information on AI markets

AI is rapidly transforming industries and economies, driving a new wave of technological innovation. Fundamentally, AI involves the development of technical systems capable of performing tasks that would otherwise require human cognitive abilities, such as language interpretation, image processing, data analysis, and decision-making.[4]

This background section is structured as follows: first, we define key AI concepts in section 2.1; second, we examine the emergence of generative AI in section 2.2; third, we map the principal layers of the AI value chain in section 2.3; and finally, in section 2.4, we highlight four critical market trends that inform the competition analysis.

2.1 Overview of AI: from narrow to general

The field of AI encompasses a range of capabilities. For clarity in discussions, it is useful to distinguish between currently deployed AI and more advanced, aspirational forms:

- Artificial Narrow Intelligence (ANI): Also known as weak AI, ANI systems are designed and trained to perform specific tasks or a defined range of specific tasks. Virtually all AI applications in use today, including the most advanced Large Language Models (LLMs), fall into this category. While LLMs, for instance, can generate human-like text, translate languages, and answer questions across a wide array of topics, their intelligence is 'narrow' in the sense that it is bound by the data they were trained on and the algorithms that govern them. They do not possess genuine 'understanding' or the ability to transfer learned knowledge into entirely novel domains outside their training in the way a human can. Their impressive capabilities stem from sophisticated pattern matching and statistical inference, not from general cognitive abilities. The consensus among most experts is that current systems, including powerful LLMs, are still far from exhibiting the robust, adaptable, and genuinely understanding intelligence characteristic of human cognition.
- Artificial General Intelligence (AGI): AGI, sometimes called strong AI, represents a theoretical future stage of AI development. An AGI system would possess cognitive abilities functionally equivalent to those of a human across a broad spectrum of intellectual tasks. This implies the capacity for true understanding, abstract reasoning, common-sense knowledge, and the ability to learn and autonomously transfer knowledge from one domain to completely unrelated domains with human-like efficiency and adaptability. AGI would not merely execute predefined tasks based on patterns but would be able to solve novel problems it was not specifically trained for, demonstrating a level of generalised intelligence and adaptability currently unique to humans. It is important to underscore that, despite rapid advancements in ANI, AGI remains a theoretical construct and has not been achieved.

[4] See, for example, IBM (2024), 'What is artificial intelligence (AI) in business?', February, https://www.ibm.com/think/topics/artificial-intelligence-business, and IBM (2023), 'Understanding the different types of artificial intelligence', October, https://www.ibm.com/think/topics/artificial-intelligence-types, accessed 28 May 2025, respectively.

- Artificial superintelligence (ASI) represents a hypothetical form of AI that surpasses human intelligence, potentially possessing cognitive capabilities—such as learning, problem-solving, and adapting to new situations—that far exceed those of humans.

While some companies claim that their models are approaching AGI, most experts contend that current systems, including those from leading developing firms, remain firmly within the realm of ANI. While these models may excel at specific tasks and demonstrate capabilities at a human-like level and sometimes beyond (e.g. tasks relating to pattern recognition), they lack the generalised cognitive abilities of AGI. When we refer to 'AI' in this paper, we are by default referring to ANI.

AI is being deployed across various sectors, demonstrating its transformative potential. For instance, in computer vision, AI empowers applications such as facial recognition, object detection in real-time environments (e.g. traffic management), and image captioning, which enhances accessibility and image search.[5] In robotics, AI enables autonomous navigation and path planning for self-driving vehicles and warehouse automation, as well as sophisticated object manipulation in manufacturing and healthcare.[6] AI is also revolutionising clinical applications, aiding early disease detection, personalised treatment, and the development of new therapies.[7]

2.2 The rise of generative AI

2.2.1 Definitions

A recent paradigm shift within AI, garnering significant global attention, is the emergence of generative AI. This prominent AI subset comprises models engineered to create novel content—including text, imagery, audio, or software code—by discerning and replicating patterns from extensive datasets. This contrasts with predictive AI (also referred to as machine learning), which primarily focuses on forecasting future outcomes or classifying events based on historical data.[8]

Current generative AI models are built on neural networks—a computational architecture that uses a layered, interconnected processing structure. These models undergo training on massive and diverse datasets, equipping them with a broad base of capabilities adaptable to a wide array of downstream applications. Due to their inherent adaptability and cross-domain flexibility, these models are referred to as Foundation Models (FMs) as they enable efficient transfer learning to new problems and reduce the necessity for extensive, task-specific

[5] AWS, 'Amazon Computer Vision', accessed 28 May 2025.
[6] Zhang, J. (2021), 'AI based Algorithms of Path Planning, Navigation and Control for Mobile Ground Robots and UAVs', October, https://arxiv.org/abs/2110.00910, accessed 28 May 2025.
[7] Vilhekar, R.S. and Rawekar, A. (2024), 'Artificial Intelligence in Genetics', *Cureus*, **16**:1, January, https://pmc.ncbi.nlm.nih.gov/articles/PMC10856672/, accessed 28 May 2025.
[8] The distinction between AI and machine learning is not always clear cut as, strictly speaking, machine learning is widely considered to be a subset of AI and current AI is based on machine learning algorithms. Nonetheless, this is a helpful distinction for discussion.

retraining. Subsequent fine-tuning on smaller, specialised datasets further refines their performance for specific use cases.[9]

LLMs are a prominent type of generative AI, trained on text data (e.g. OpenAI's GPT, Google's Gemini, or Anthropic's Claude). Increasingly, we also see multimodal models, which can handle several types of data, such as text and images concurrently. However, LLMs are currently the most prevalent type of generative AI model in the market.[10] Box 2.1 below introduces some of the most important technical terms in generative AI that will be used throughout this briefing paper.

Box 2.1 Select technical concepts in AI models

Parameter: An internal variable within an AI model, adjusted during the training phase to optimise performance and enable learning from data. Parameters function as the model's 'knobs', refined based on input data to adapt to complex patterns and make accurate predictions.

Token: in LLMs is the smallest unit of text that a model can process. Tokens can be individual words, sub-parts of words, or even punctuation marks, depending on the model's tokenisation strategy.

Context window: in LLMs, refers to the amount of information, typically measured in tokens, that the model can consider at any given time. It is analogous to a system's working memory, determining how long a conversation can be had before the model begins to lose recall of earlier details.

Transformer architecture: is a multi-layer neural network architecture (based on 'deep learning') introduced by Google researchers in the 2017 paper 'Attention is All You Need.'[11] The Transformer architecture underpins many state-of-the-art generative AI models. It deviates from traditional sequential data processing by relying entirely on attention mechanisms, particularly self-attention, enabling parallel processing of input sequences for enhanced efficiency and contextual understanding.

[9] The Stanford Institute for Human-Centered Artificial Intelligence's (HAI) Center for Research on Foundation Models (CRFM) coined the term 'foundation model' in August 2021. See https://hai.stanford.edu/news/introducing-center-research-foundation-models-crfm, accessed 28 May 2025. Also, OECD (2024), 'Artificial intelligence, data and competition', OECD Artificial Intelligence Papers, 18, p. 18.
[10] Toner, H. (2023), 'What Are Generative AI, Large Language Models, and Foundation Models?', May, Georgetown University, Center for Security and Emerging Technology (CSET).
[11] Vaswani, A. et al (2017), 'Attention Is All You Need', arXiv, accessed 28 May 2025.

2.2.2 Impacts of generative AI in the economy

AI's economic impact is likely to be profound, reshaping how economic value is created and distributed. It is increasingly recognised as a general-purpose technology (GPT), akin to electricity or the internet, possessing the potential to disrupt a vast array of sectors and economic activities. AI promises substantial productivity enhancements through task automation, augmented data analysis, and optimised resource allocation.

At the **firm level**, AI may contribute to cost reductions through automation and process optimisation across a wide range of specific industries and applications, particularly in areas such as logistics, customer service, and manufacturing. Furthermore, AI enables the creation of innovative products and services previously unfeasible due to technical or economic limitations, such as highly personalised digital experiences. In addition, AI facilitates the development of products and services that were previously unviable or unimagined due to technical or economic constraints, such as highly personalised digital services.

Illustrative use cases for generative AI include:[12]

- **Customer service:** AI empowers businesses to deliver more personalised and efficient customer support through sophisticated chatbots, virtual assistants, and AI-augmented tools for sales, marketing, and human agent support, providing real-time insights to improve interaction quality and customer satisfaction.
- **Talent acquisition:** AI is streamlining key human resources functions, including automated candidate sourcing and pre-screening, optimised onboarding processes, and enhanced compliance management across diverse jurisdictions.
- **Operational automation:** AI tools automate repetitive operational tasks, enabling employees to concentrate on higher-value strategic activities. This includes coding assistants, backend process automation for internal workflow streamlining, and no-code platforms for rapid digitisation.

Other business uses for multimodal models include image and audio generation. For example, they can be used to generate marketing materials, product mock-ups, 3D models, or audio for commercials. Midjourney and DALL-E 3 are among the most popular AI image generators, the latter also known for its integration with ChatGPT.[13]

The **labour market** effects of AI are multifaceted. Occupations involving routine or highly codifiable tasks face potential automation, possibly leading to job displacement or wage stagnation. The effects on other tasks are ambiguous, as some professions might be vulnerable (e.g. photography or writing), however experience increased demand nonetheless, either through direct complementarity with AI systems or by shifting toward tasks that remain difficult to automate, such as interpersonal interaction or complex decision-making.

Taken together, these developments illustrate potentially significant impacts of AI worldwide. Goldman Sachs estimates that generative AI could drive a 7% (or almost $7trn) increase in

[12] Puutio, A. (2024), 'Real-World Uses Of AI In Business: What You Need To Know Today', Forbes, 20 September, accessed 28 May 2025.
[13] Foley, J. (2024), 'The best AI image generators', 22 October, accessed 28 May 2025.

global GDP and lift productivity growth by 1.5 percentage points over a ten-year period.[14] Their research also indicates that approximately two-thirds of US occupations are exposed to some degree of AI automation, with between one-quarter and one-half of the workload in these occupations being potentially replaceable. However, most jobs and industries exhibit partial exposure, suggesting AI will more likely complement rather than fully substitute human labour.

While the full extent of AI's economic impact remains subject to ongoing assessment, its expanding role across markets is undeniable and is poised to reshape competitive landscapes.

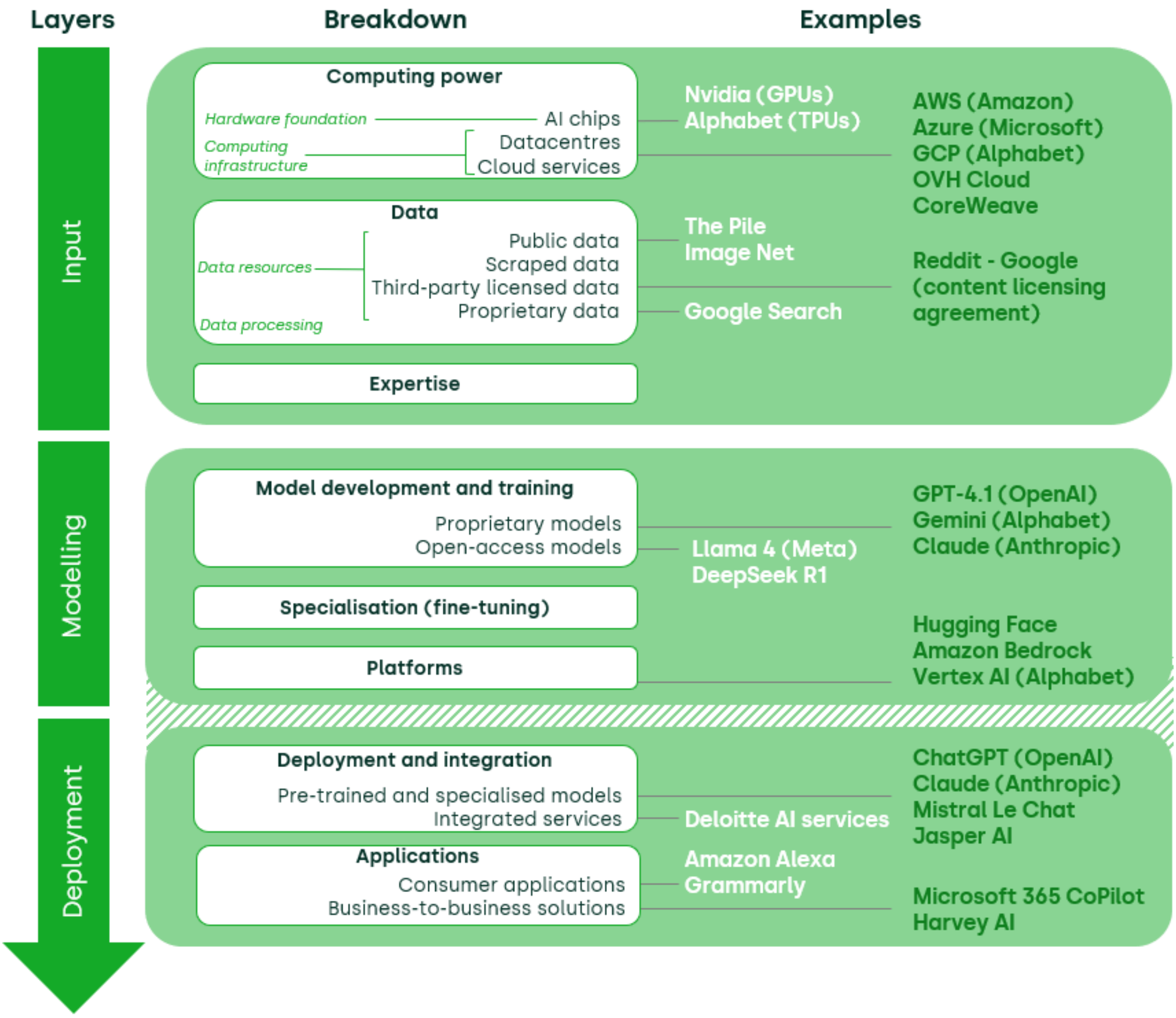
2.3 The generative AI value chain

The generative AI ecosystem can be conceptualised through a value chain comprising three layers, as depicted in Figure 2.1 below. Each layer constitutes a distinct phase throughout the lifecycle of a generative AI model.

- The **input** layer is where the fundamental components of AI—computing power, data, and expertise—are sourced. These elements are the building blocks for creating general-capability AI models.
- The **modelling** layer is where models at the core of generative AI are developed and trained.
- The **deployment** layer is where these models are implemented and delivered as products or services to end-users (consumers, businesses, or public sector entities).

[14] Goldman Sachs (2023), 'Generative AI could raise global GDP by 7%', 5 April, accessed 28 May 2025.

Figure 2.1 Generative AI value chain



Source: Oxera.

While these layers provide a useful conceptual framework, reality is more intricate. Adjacent or intermediate layers exist, such as platforms for model management (see section 2.3.2). Furthermore, the distinction between development and deployment can be blurred when models are continuously fine-tuned or adapted for end-uses. The generative AI landscape is also characterised by vertically integrated companies often operating across multiple layers, rendering the value chain more complex.

The following sections provide a more detailed view of the three layers in our value chain framework.

2.3.1 The input layer

The development of a generative AI model necessitates three primary components: computing power, data, and a skilled workforce (expertise).

Computing power

Computing power refers to 'all services offering the capacity to process large amounts of information at the same time'. [15] Computing power enables the execution of training algorithms that adjust model parameters based on data inputs, thereby enabling the training and operation of highly capable models. It is generally recognised in the industry that access to substantial computing hardware is expensive, compounded by energy consumption (e.g. electricity and cooling), significant space requirements,[16] and costly datacentre maintenance.

Computing capacity requires high-performance IT hardware, generally via thousands of special-purpose chips, such as **graphics processing units** (GPUs).[17] These chips are capable of performing a large number of operations in parallel. To access computing power, industry players have the choice to either develop on-site infrastructure (i.e. build their own datacentres), use cloud computing, or more rarely, use public supercomputers as a shared resource.[18] For companies building their own datacentres (including cloud service providers (CSPs)), access to the relevant hardware means either purchasing AI chips from specialised GPU designers such as Nvidia, or designing their own special-purpose chips and having them produced by a third party.

AI chips

The sector of AI chips is highly concentrated but is experiencing entry from capable, well-resourced new entrants. The clear leading producer of commercially available AI chips is American company Nvidia, which sells over 90% of datacentre GPUs worldwide. Nvidia has maintained a strong position in the sector in recent years. However, competitors such as AMD and Intel remain active, and the industry is also seeing the emergence of new players.[19]

Cloud providers such as Amazon, Google and Microsoft have started to develop their own, purpose-built, AI chips in order to lower their dependence on Nvidia. In 2018, Google made its Tensor Processing Units (TPUs) available to its cloud customers, tailored for specialised AI use

[15] ADLC (2023), 'opinion 23-A-08 on competition in the cloud sector', 29 June, para. 28.
[16] Stakeholder's response to questionnaire, ADLC (2024), 'opinion 24-A-05 on the competitive functioning of the generative artificial intelligence sector', 28 June, para. 130.
[17] OECD (2024), 'Artificial intelligence, data and competition', OECD Artificial Intelligence Papers, No. 18, OECD Publishing, Paris. See also ADLC (2024), 'opinion 24-A-05 on the competitive functioning of the generative artificial intelligence sector', 28 June.
[18] For example, a team of researchers from the CentraleSupélec University has trained a model called 'CroissantLLM' on the French supercomputer Jean Zay. See ADLC (2024), 'opinion 24-A-05 on the competitive functioning of the generative artificial intelligence sector', 28 June, p. 5.
[19] IOT Analytics (2025), 'The leading generative AI companies', March.

cases.[20] Similarly, Amazon made its Inferentia and Trainium 2 AI solution available to its Amazon Web Services (AWS) customers in 2023. Microsoft followed in 2024 with the release of its Maia AI accelerators to Azure customers. Beyond cloud providers, large tech companies such as Apple and Meta are also investing in AI-specific IT hardware.[21,22]

Despite designing AI chips, companies such as Nvidia and the other large tech firms do not operate their own fabrication plants ('fabs').[23] Rather, chip designs are manufactured by foundries, with Taiwan's TSMC being the predominant producer for high-end AI chips. To support and enable production, TSMC procures specialised machines from the Dutch company ASML. Companies such as Nvidia, Intel, Microsoft, and Alphabet rely on TSMC for their AI chip orders, indicating that the upstream value chain is currently highly concentrated and prone to supply constraints.[24]

The semiconductor industry has faced significant shortages, exacerbated by demand surges and supply chain disruptions. Waiting times for high-demand chips such as Nvidia's H100, which reached nearly 12 months at the end of 2023, moderated to three to four months in 2024.[25] Prices remain generally high (reported to be up to 30 or 40 thousand US dollars per unit, and potentially higher for leading-edge versions),[26] although more accessible alternatives are emerging.[27] The recent surge in demand for AI across all industries could cause a subsequent world chip shortage,[28] although innovations including the increasing efficiency of models may offer some mitigation.[29]

Cloud computing

The rapid obsolescence of AI hardware, particularly for generative AI, makes on-premise infrastructure a challenging proposition.[30] Scaling with the latest hardware requires constant

[20] Tarasov, K. (2024), 'How Google makes custom chips used to train Apple AI models and its own chatbot, Gemini', Cnbc.com, 23 August, accessed 28 May 2025. TPUs are not widely available to the market. To the contrary, Nvidia or AMD's GPUs are available in various options and give users flexibility in their computing set up. GPUs are more popular in a range of industries whereas TPUs are more common in cloud-focused areas and sectors such as deep learning. See datacamp.com (2024), 'Understanding TPUs vs GPUs in AI: a comprehensive guide', 30 May, accessed 28 May 2025.
[21] Paul, K. and Hu, K. (2025), 'Meta begins testing its first in-house AI training chip', Reuters, 11 March, accessed 28 May 2025.
[22] Ankursnewsletter.com (2024), 'Google TPUs vs. AWS Trainium & Inferentia vs. NVIDIA GPUs', 7 November, accessed 28 May 2025.
[23] The upstream segment of silicon production is outside the scope of this paper.
[24] Unconfirmed recent reports suggest that Nvidia may turn to Samsung's foundry (TSMC's trailing competitor) for producing less powerful AI processors, diversifying AI chip orders with TSMC. Google is also currently partnering with Samsung Foundry but reports to turn to TSMC for the next generation of chipsets. See: Pol, F.H. (2025), 'Samsung may finally supply Nvidia HBM3E memory chips', Techzine, 30 January; Garreffa, A. (2025), 'Nvidia could have 2nm AI GPUs fabbed at Samsung Foundry', TweakTown, 12 May; TrendForce (2024) 'TSMC Rumored to secure Tensor G6 orders from Google with its 2nm Process', 12 September, accessed 28 May 2025 respectively.
[25] TrendForce (2024), 'NVIDIA's H100 AI Chip No Longer Out of Reach, Inventory Pressure Reportedly Forces Customers to Resell', accessed 29 May 2025.
[26] Sourcengine (2025), 'Semiconductor Industry News - January 2025 Update', accessed 29 May 2025.
[27] Mo, L. and Potkin, F. (2025), 'Exclusive: Nvidia to launch cheaper Blackwell AI chip for China after US export curbs, sources say', Reuters, accessed 29 May 2025.
[28] Kharpal, A. (2024), 'Surging AI demand could cause the world's subsequent chip shortage, research says', Cnbc.com 25 September, accessed 28 May 2025.
[29] For further developments on the increasing model efficiency and the example of DeepSeek's efficiency breakthrough, see section 2.4.2 and Box 2.2 respectively.
[30] ADLC (2024), 'opinion 24-A-05 on the competitive functioning of the generative artificial intelligence sector', 28 June.

reinvestment. Cloud providers benefit from recurring cash flows by serving multiple customers, allowing them to spread costs and risks, unlike AI developing firms who face uncertain visibility into future hardware needs. Consequently, cloud computing (for which the obsolescence burden lies on the cloud provider) has emerged as a widely adopted solution for training and fine-tuning models,[31] offering a way for AI-developing firms to minimise maintenance overheads while gaining flexibility and scalability, albeit at a significant cost.[32]

The cloud computing sector is highly concentrated, especially for IaaS (Infrastructure as a Service) and PaaS (Platform as a Service),[33] with AWS, Microsoft Azure and Google Cloud Platform (GCP) acting as the main providers.[34] Due to their large storage and computing capacities, the three companies are typically referred to as 'hyperscalers' in cloud computing,[35] and coexist with other much smaller 'pure players' which provide only cloud computing services, such as 3DS Outscale, OVHCloud, Clever Cloud or Qarnot Computing. Specialised GPU-cloud providers such as CoreWeave have also successfully entered the market and seen rapid growth. Worldwide, hyperscalers collectively account for 63% of global cloud spending, including 30% attributable to Amazon (AWS).[36] Hyperscalers seem to have maintained their position in the industry over the last years, with slight variations in market shares.[37]

Data

Data is the lifeblood of AI models; leading FMs, in particular, are trained on vast datasets. Procuring data is an important aspect of the AI value chain, but equally important is data processing, cleaning the data to remove errors or inconsistencies, and labelling it so the model can learn from it.

According to stakeholders, the majority of generative AI models are trained, at least in part, using publicly accessible data,[38] some of which is acquired through web scraping and crawling. This includes datasets such as (for text) the Pile and the Common Crawl, (for images) ImageNet and Open Images, and (for audio) LibriSpeech, but also books, news articles, scientific journals, forums, user-generated content on social media, etc.[39] Many firms have not disclosed the specifics of their training datasets beyond stating that they use a

[31] Ibid.
[32] Insiders estimate that, on average, AI application developers 'spend around 20-40% of revenue on inference and per-customer fine-tuning […] typically paid either directly to cloud providers for compute instances or to third-party model providers – who, in turn, spend about half their revenue on cloud infrastructure'. Kinsella, B. (2023), 'The most interesting analysis of the Generative AI market to date has arrived, Synthedia, 20 January, accessed 28 May 2025.
[33] IaaS provides computing resources like servers and storage where the user can manage the software. PaaS offers ready-to-use platforms with tools for developing and deploying apps. At the infrastructure layer, CSPs offer IaaS services.
[34] The CMA recently found that AWS and Microsoft together hold a market share in the UK of between 70% and 90%. Longo, R. and Rocha, M. (2024), 'Generate AI: The new digital frontier for competition', *Concurrences no.2-2024*, p. 24. CMA (2025), 'Cloud services market investigation', summary of provisional decision, para. 6.
[35] ADLC (2023), 'opinion 23-A-08 on competition in the cloud sector', 29 June, p. 4.
[36] Richter, F. (2025), 'Amazon and Microsoft Stay Ahead in Global Cloud Market', Statista.com, 27 February, accessed 28 May 2025.
[37] ADLC (2023), 'opinion 23-A-08 on competition in the cloud sector', 29 June.
[38] ADLC (2024), 'opinion 24-A-05 on the competitive functioning of the generative artificial intelligence sector', 28 June, para. 149.
[39] Alaily, M. (2024), 'The new AI economy: understanding the technology, competition, and impact for societal good', *Concurrences no.2-2024*, p. 5.

mixture of public and licensed data.[40] It has been repeatedly suspected that the training data of leading models contains significant amounts of copyrighted material without having obtained an explicit licence for this (whether and where such a licence would be needed is at the centre of several ongoing legal disputes, see Section 3).

Uncertainties remain regarding the future demand for new, proprietary, or specialised datasets, as further discussed in section 3.2.2.[41] There is a growing concern that original data sources will soon be fully exploited, leaving a dearth of novel data for training future models.[42] Although small quantities of data can be extrapolated to generate synthetic data, the viability of such data replacing real, human-created data is a subject of ongoing debate. The current consensus suggests that synthetic data is unlikely to generate sufficient value for training innovative FMs, as research indicates that models trained predominantly on synthetic data can suffer from 'model collapse'—a degenerative process where, over time, the models lose fidelity.[43]

Having access to large volumes of high-quality data plays an instrumental role in the accuracy of generative AI outputs.[44] Publicly accessible data may be suboptimal or insufficient in this regard, as high-quality, curated datasets appear to be crucial for generating high-quality output.

Digital conglomerates such as Alphabet, Microsoft, Amazon, or Meta, with their extensive user bases and existing data repositories, are potentially powerful incumbents in data provision. The potential effects associated with these major digital companies having preferential access to proprietary data, and the broader implications of their vertical integration across the AI value chain, are explored further in sections 3.2.2 and 3.2.3.

Expertise

The development and efficient deployment of AI models demand a highly skilled workforce. Advanced capabilities in data science, machine learning, deep learning, natural language processing (NLP), computer vision, software engineering, and development operations (DevOps) are indispensable for architecting and optimising the complex systems that run AI models.[45] The high demand for such specialised skills raises concerns about talent availability.

However, anxieties about short-term access to expertise may be partially mitigated. The French competition authority (ADLC), for example, notes the emergence of successful start-

[40] OECD (2024), 'Artificial intelligence, data and competition', OECD Artificial Intelligence Papers, 18.
[41] Kowalski, K., Volpin, C. and Zombori, Z. (2024), 'Competition in Generative AI and Virtual Worlds', *Competition Policy Brief*, 3, September.
[42] CMA (2023), 'AI foundation models', Short Version.
[43] Hunt, S., Jian, W., Mawar, A. and Tablante, B. (2023), 'You Are What You Eat: Nurturing Data Markets to Sustain Healthy Generative AI Innovation', *CPI Tech Reg Chronicle*, Vol. 1.
[44] Quality of data depends on volume, velocity, variety and veracity (or value) of the data. Together these are known as the four V's of data. See OECD (2020), 'Consumer Data Rights and Competition', background note, June, pp. 10–11.
[45] Kowalski, K., Volpin, C. and Zombori, Z. (2024), 'Competition in Generative AI and Virtual Worlds', *Competition Policy Brief*, 3, September. See also ADLC (2024), 'opinion 24-A-05 on the competitive functioning of the generative artificial intelligence sector', 28 June.

ups, such as Mistral AI, built around a small nucleus of exceptionally skilled individuals.[46] Moreover, as highlighted by the OECD and academic research, the recent surge in demand for AI expertise has drawn considerable talent into the industry, suggesting that the labour supply is likely to expand in the longer term,[47] a typical adjustment driven by market forces.

AI experts increasingly exhibit global mobility and are highly sought after.[48] This talent circulation between established corporations, research institutions, and start-ups is vital for knowledge dissemination and fostering a vibrant ecosystem. However, it also presents challenges: larger, well-resourced firms possess considerable capacity to attract and retain top talent (e.g. through 'acqui-hires', vast resources, and generous compensation packages), potentially hindering the growth of smaller competitors.[49]

2.3.2 The modelling layer

This layer focuses on the development, training, and fine-tuning of AI models, with AI development firms at its core. Generative AI models are built by training a machine learning algorithm on a large dataset, in order to produce a model that can be refined and used in several downstream applications—for instance, 'general-purpose, grammatically correct, and contextually coherent text output'.[50] Most of the latest-generation models are built on transformer architectures, adept at capturing contextual relationships in sequential data such as text. The field is exceptionally dynamic, with continuous advancements in model architecture, training methodologies, and performance optimisation.[51]

Once trained, a model can be further specialised through fine-tuning: adapting an existing model to a specific task or domain by training it on a smaller, more relevant dataset.[52] Rather than building from scratch, many developers and researchers opt to fine-tune available pre-trained models. This approach leverages the knowledge embedded in the original model, enabling high performance on specialised tasks while requiring far less data and computing power for training.

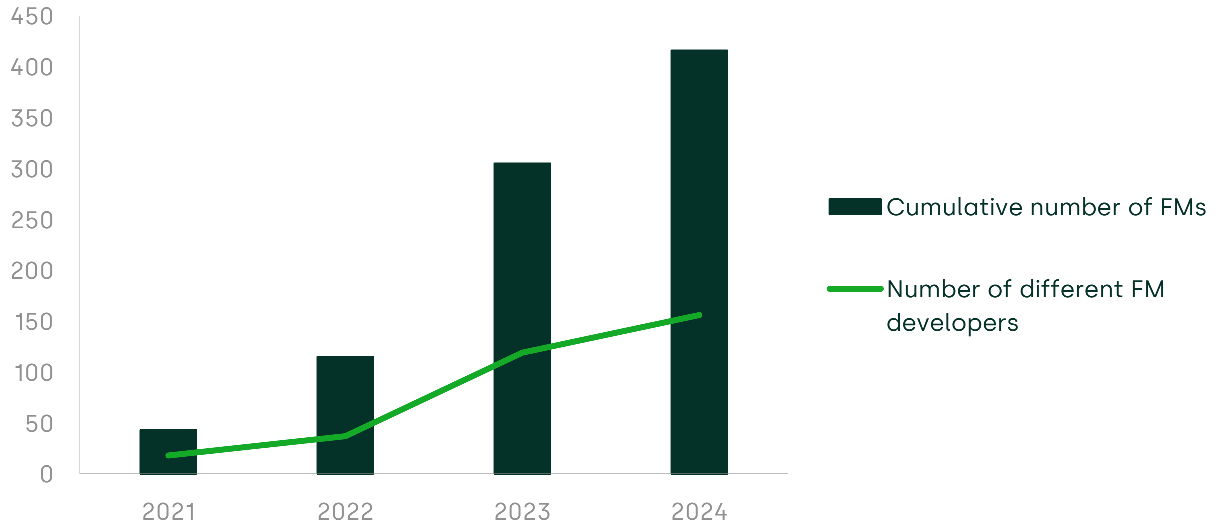
[46] ADLC (2024), 'opinion 24-A-05 on the competitive functioning of the generative artificial intelligence sector', 28 June.
[47] OECD (2024), 'Artificial intelligence, data and competition', OECD Artificial Intelligence Papers, 18, p. 35; Stanford University (2024), 'Artificial Intelligence Index Report 2024', Human-Centered AI.
[48] Brazier, J. (2024), 'AI talent is becoming more mobile and seeking new global destinations for work: BCG', Unleash, 9 December, accessed 28 May 2025.
[49] CMA (2023), 'AI Foundation Models Initial Report', initial report, 18 September, para. 3.12.
[50] Hunt, S., Jian, W., Mawar, A. and Tablante, B. (2023), 'You Are What You Eat: Nurturing Data Markets to Sustain Healthy Generative AI Innovation', *CPI Tech Reg Chronicle*, Vol. 1.
[51] For example, researchers have recently developed a multiscale decoder architecture that outperforms traditional transformers on high-dimensional outputs, such as images or lengthy text sequences, which could significantly boost the performance of FMs in the near future. See Yu et al. (2023), 'MEGABYTE: Predicting Million-byte Sequences with Multiscale Transformers', arXiv, May, https://arxiv.org/pdf/2305.07185.pdf.
[52] Alaily, M. (2024), 'The new AI economy: understanding the technology, competition, and impact for societal good', *Concurrences no.2-2024*, p. 5. Rather than building from scratch, many developers and researchers opt to fine-tune publicly available pre-trained models, significantly reducing both time and cost. This approach leverages the knowledge embedded in the original model, enabling high performance on specialised tasks while requiring far less data and computing power.

The modelling layer of the AI value chain is characterised by vibrant competition, with numerous emerging entrants and rapid innovation.[53]

Model developing companies

The number of generative AI models is increasing, with hundreds of different FMs, as well as a vast array of emerging players each year, as shown in Figure 2.2. A selected overview of generative AI model developing firms is presented in Table 2.1 (see below in section 2.3.3).

Figure 2.2 Yearly number of FMs and of FM developing firms since 2021



Note: The count of FMs is based only on models that are in the public domain. Partnerships between two organisations are counted as a separate model developer.
Source: Oxera based on Stanford CRFM (2025), 'Ecosystem Graphs for Foundation Models'.

Large digital companies such as Alphabet, Microsoft and Amazon are active across the entire value chain. They operate at the input level, owning cloud platforms and datacentres, and have developed their own proprietary models, such as Google's Gemini, Microsoft's Prometheus, Amazon's Nova, and Meta's LLaMA models.

In parallel, several successful start-ups have emerged in model development. OpenAI's GPT models and Anthropic's Claude are leading examples of high-performing closed-source (proprietary) systems. Meta, DeepSeek, and Mistral AI, on the other hand, have released models which are available under open licences ('open-weights'), offering more flexible, cost-effective alternatives and pushing for a more open AI ecosystem (see Table 2.1 below). As

[53] Alaily, M. (2024), 'The new AI economy: understanding the technology, competition, and impact for societal good', *Concurrences no.2-2024*, p. 9.

discussed above, there is an additional number of emerging players that build their AI models by fine-tuning existing ones (e.g. Tenyx fine-tuning Meta's LLaMA to create Tenyx-70B).[54]

Model management platforms

At the intersection of the modelling and deployment layers, model management platforms have emerged as software solutions that enable users to train, fine-tune, and deploy AI models. For example, companies like Hugging Face act as hubs for open-source models and tools. Large digital companies also provide such platforms, typically through cloud-based PaaS offerings. Notable examples include Microsoft (Azure AI), Amazon (Amazon Bedrock), and Alphabet (Vertex AI).

2.3.3 The deployment layer

The third layer of the value chain involves the deployment and application development of AI products and services for end-users. This layer is particularly relevant when end-users, including businesses, wish to use AI as an input in their products and services.

AI products or services can be delivered to end-users (consumers, public bodies or businesses) in various forms: as an *ad hoc*, in-house developed AI model (for example, Chat-GPT); as an application that integrates one, or numerous, AI models (for example, Grammarly[55]); or as a service solution composed of multiple models. For example, in the latter category, Jasper's AI Engine (a generative AI platform designed primarily for marketing, content creation, and enterprise use) works using various custom, proprietary, in-house models as well as several third-party models.[56] Likewise, large IT services firms including Accenture, IBM, and others deliver AI solutions that combine proprietary tools, customised AI models, and third-party components, tailored to specific needs and integrated into existing business processes.

Demand for AI products and services is strong and increasing—from consumer applications to industry solutions. Companies have rapidly integrated AI tools into their products (see section 2.2.2), with the proliferation of a wide range of chatbots or personalisation experiences, across a range of industries from AdTech[57] to healthcare[58] or online shopping.[59] Associated with an emerging number of AI developing firms, and the various forms under which AI products and services can be offered, supply for AI products and services has become highly fragmented with numerous suppliers operating in this layer.

[54] See Nunez, M.F. (2024), 'Exclusive : AI startup Tenyx's fine-tuned open-source Llama 3 model outperforms GPT-4', VentureBeat, 7 May, accessed 28 May 2025.
[55] English language writing assistant software tool.
[56] See Jasper, 'The Jasper AI Engine' accessed 28 May 2025.
[57] Yoo, H. (2024), 'Meet the Ad Tech Players Using Generative AI For Their Media Buys', AdExchanger, 13 March, accessed 28 May 2025.
[58] Insilico Medicine, https://insilico.com/, accessed 28 May 2025.
[59] Deremuk, I. (2025), 'Artificial Intelligence Examples: How Alibaba, Amazon, and Others Use AI', LITSLINK, 5 May, accessed 28 May 2025.

Table 2.1 presents a non-exhaustive list of generative AI models deployed for end-use. The available information on the count of weekly users indicates that ChatGPT is the most widely used AI tool worldwide, followed by Alphabet's Gemini model.

Table 2.1 Select leading LLMs in May 2025

Deployed model	Developer	LLM name	Access	Users
ChatGPT	OpenAI	GPT-o4-mini, GPT-4.1	Closed source	400m weekly
Gemini	Alphabet	Gemini	Closed source	285m weekly
LLaMA	Meta	LLaMA 3	Open under restricted conditions[60]	Unknown
Quark	Alibaba	Qwen 3	Closed source	[38–150]m weekly
Qwen	Alibaba	Qwen 2.5-Max	Open under licence	Unknown
CoPilot	Microsoft, OpenAI	Prometheus, GPT-4, GPT-4 Turbo	Closed source	20m weekly
DeepSeek	DeepSeek	DeepSeek R1	Open under licence	[15–60]m weekly
Grok	xAI	Grok 3	Open under licence	[9-35]m weekly
Claude	Anthropic	Claude Family	Closed source	[5–19]m weekly
Mistral AI	Mistral AI	Mistral Large 2, Mixtral 8x22B	Open under licence	[2–8]m weekly

Note: The range represents the monthly users divided by four to approximate the lower bound of weekly users, while the monthly number of users is provided for the upper bound.
Sources: Guadamuz, A. (2025), 'How many people are using generative AI on a daily basis? A Gemini report', TechnoLlama, 14 April; Campbell, I.C. (2025), 'It seems like most Windows users don't care for Copilot', engadget.com, 25 April; 'DeepSeek License FAQ', https://deepseeklicense.github.io/; Backlinko (2025), 'DeepSeek AI Usage Stats' 27 May; and Backlinko (2025), 'Claude Statistics: How Many People Use Claude?', 3 March; Cardillo, A. (2025) 'Number of Grok users' (accessed 28 May 2025, respectively).

2.4 Trends in AI development

Recent developments in AI markets reflect a rapidly evolving technological landscape shaped by three interrelated trends: (i) the parallel advancement of open-weight and proprietary models; (ii) the bifurcating size of models; and (iii) the growing interconnectedness across platforms and applications. These dynamics are not only reshaping the structure of the AI ecosystem; they are also raising important questions for regulation. Understanding these trends is essential for assessing policy interventions.

[60] A commercial use licence should be requested for companies with over 700m monthly active users. See Maris, J. (2025), 'Meta's LLaMa license is still not Open Source', Open Source Initiative, 18 February, accessed 28 May 2025.

2.4.1 Parallel development of open source and proprietary models

While many of the leading LLMs remain proprietary, there has been rapid and significant innovation within the open-model ecosystem. Closed models are ones where neither the weights nor the model are accessible to the public. In contrast, fully open models are ones where all the code, architecture, training data, weights and the learning process are made available to the general public.[61] Most non-closed models available on the market are not fully open. Rather, these are 'open-weights' models where only the model itself is publicly made available, while the training data, architecture, or source code used for training remain undisclosed (for instance, LLaMA, see Table 2.1 above).

A notable example of an efficient open-weights model is Alpaca, an LLM developed by researchers at Stanford University. Based on Meta's LLaMA 7B model, Alpaca is fine-tuned using 52,000 instruction-following samples generated by OpenAI's text-davinci-003 model. The fine-tuning process cost under $600[62] and the result is an improved performance in tasks such as text generation, summarisation, and conversational AI. Alpaca can generate responses similar to GPT-based models and is freely accessible to developers and researchers. However, as observed in Table 2.1, most popular AI services are still based on proprietary models developed by large firms.

2.4.2 Bifurcating size of models

Since OpenAI introduced GPT-1 in 2018, each new GPT model has been larger than the previous one. GPT-3 with 175bn parameters is more than 100 times larger than GPT-2 with 1.5bn parameters, and GPT-4 is even larger than GPT-3.[63]

Context window size—the maximum number of tokens that the model can process at one time—has increased from 1,024 tokens in GPT-2 to 2,048 tokens in GPT-3, allowing the model to handle longer input sequences. The context window in GPT-4 is larger than in GPT-3, up to 32,000 tokens in some versions. Similarly, Anthropic's Claude 3 has a context window size of 200,000 tokens.[64] Google stated that its Gemini 1.5 Pro is available with a context window of up to one million tokens (over 700 thousand words or one hour of video).[65]

On the other hand, the LLM landscape is increasingly shaped by the proliferation of smaller, more efficient, and less data-intensive models. While some launches have attracted more public attention than others (e.g. see Box 2.2 below, on DeepSeek),[66] the pursuit of compact, cost-effective models appears just as intense as the pursuit for ever larger leading-edge

[61] ADLC (2024), 'opinion 24-A-05 on the competitive functioning of the generative artificial intelligence sector', 28 June, para. 179.
[62] Taori, R. et al. (2023), 'Alpaca: A Strong, Replicable Instruction-Following Model', Stanford CRFM, March, accessed 28 May 2025.
[63] Koti, V. (2024), 'Evolution of GPT Models, GPT 1 to GPT 4', Medium, 12 September, accessed 28 May 2025.
[64] Anthropic, 'Model comparison table', accessed 28 May 2025.
[65] Pichai, S. and Hassabis, D. (2024), 'Our next-generation model: Gemini 1.5', Google, 15 February, accessed 28 May 2025.
[66] Martin, H. (2025), 'DeepSeek: Chaos Hits Western Tech as '$6M' Chinese AI Emerges', Salesforce Ben, accessed 28 May 2025.

models.[67] Much of the progress in efficiency—beyond hardware improvements—has come from advances in LLM architecture and training data strategies.[68]

Box 2.2 DeepSeek efficiency breakthrough

DeepSeek has developed a number of LLMs that provide responses comparable to other high-performance LLMs—such as OpenAI's o1—with only a fraction of the usual computational cost. Seeking an efficiency-driven AI design, developers trained the models on fewer high-quality instruction samples (sourced from larger proprietary models) than what would typically be used for models of such performance levels.

For instance, in terms of input and output costs for processing tokens in user prompts, DeepSeek-V3 shows to be over 200 times cheaper than Open AI's GPT-4. As one of the firm's latest LLM iterations, DeepSeek-V3 was supposedly trained in approximately two months, for only USD $5.6m, which amounts to roughly 1% of the budget of comparable projects. Similar cost-efficiencies can be identified when comparing R1, DeepSeek's latest LLM, to OpenAI's latest model o3.

By making these highly efficient models openly available, DeepSeek has achieved a broad adoption of their technologies. In February 2025, DeepSeek had over 60m monthly active users worldwide, and recent trends suggest that their user base will continue to grow significantly.

Source: Oxera analysis based on DocsBot AI (2024), 'DeepSeek-V3 vs GPT-4 - Detailed Performance & Feature Comparison', accessed 20 May 2025; Backlinko (2025), 'DeepSeek AI Usage Stats', accessed 21 May 2025; Jiang, B. and Perez, B. (2025), 'Meet DeepSeek: the Chinese start-up that is changing how AI models are trained', South China Morning Post, accessed 21 May 2025; DocsBot AI (2025), 'DeepSeek-R1 vs o3 - Detailed Performance & Feature Comparison', accessed 21 May 2025.

For example, the s1 LLM, developed by researchers at Stanford University and the University of Washington, reportedly uses only 16 GPUs and 1,000 data points while matching the performance of leading models like OpenAI's o1 and DeepSeek-R1, at a training cost of just

[67] Soni, A. and Kchwala, Z. (2025), 'DeepSeek's low-cost AI spotlights billions spent by US tech', Reuters, accessed 28 May 2025.
[68] This is among many other developments in AI. For example, 'reasoning' LLMs (e.g. OpenAI's GPT-o1) break down a user's question into smaller parts, plan how to answer it, and then refine their response step-by-step. This helps ensure that the answer is clear and on target. Another example is using Retrieval-Augmented Generation (RAG) architecture designed to mitigate hallucinations by incorporating real-time internet validation ('Deep Research' models by OpenAI and Perplexity). See, for example: Gillespie, T. (2025), '[AI SPRINT] What Is Deep Research AI? Comparing OpenAI, Google, Perplexity & X.AI, AI SPRINT, 19 February; Liu, J., Lin, J. and Liu, Y. (2024), 'How Much Can RAG Help the Reasoning of LLM?', arXiv; and Google, 'What is Retrieval-Augmented Generation (RAG)?' (accessed 28 May 2025, respectively).

$50. However, it is worth noting that these 1,000 data points were carefully selected from a larger 59,000-point dataset to optimise training time.[69]

Recent benchmarking exercises consistently reveal a pluralistic landscape: no single model dominates across every task. Instead, each architecture brings its own mix of strengths and trade-offs.

2.4.3 Tendency towards interconnectedness

Beyond rapid technological advancements, the AI sector is characterised by increasing vertical integration, partnerships, and strategic agreements, creating an ever more interconnected market. Competition authorities, including the UK's CMA, have highlighted this ‘interconnected web’ of over 90 partnerships and strategic investments involving key players such as Alphabet, Apple, Microsoft, Meta, Amazon, and Nvidia.[70] These collaborations frequently feature:

- **Investment, equity and revenue-sharing agreements**: the partnerships may provide CSP partners with a share of the equity and/or a right to a part of the profits of the AI developer partners (e.g. Amazon/Anthropic).
- **Expertise sharing agreements**: the partners agree to share their expertise. The partnerships may provide the CSP partners with consultation rights. They may also provide for information transfer by embedding engineers in one another’s companies, or by giving access to IP rights (e.g. Microsoft/Inflection AI or Google/CharacterAI).
- **Data sharing agreements**: the partnerships may facilitate data licensing agreements (e.g. OpenAI/Le Monde).
- **Compute sharing agreements**: the CSP partner may give access to compute infrastructure to the AI developer partner. This may include access to cloud services and co-development plans for AI-optimised CSP-designed semiconductor chips (e.g. Microsoft/OpenAI, Amazon/Anthropic, and Microsoft/MistralAI).
- **Distribution partnerships**: partnerships focused on the deployment of AI models across products and services (e.g. Microsoft/Meta, Microsoft/Nvidia, etc.).

The CMA recognises that these connections may have pro-competitive effects as it allows complementary products and capabilities to be utilised to create innovative products. They can be important to maintain competition at the development level of the AI value chain as they can be especially important for smaller firms which may otherwise struggle to compete in FM development and related service markets. Smaller firms enjoy the significant efficiencies that these strategic partnerships can offer—these usually include easier knowledge sharing, reduced costs, enhanced product quality, more efficient resource allocation, more direct access to customers and consumers,[71] and access to the necessary computing power.

[69] The Economist (2025), ‘Forget DeepSeek. Large language models are getting cheaper still’, 12 February.
[70] CMA (2024), ‘AI Foundation Models: Update paper’, 11 April, p. 10.
[71] For example, partnerships can include integration of AI FM functionalities into an established product of a larger digital player, or more targeted access to customers or consumers.

Access to compute power is indeed important to market players; AI developing firms often rely on partnerships with established tech firms to secure critical sufficient infrastructure needs, which can help them scale and innovate (see, for example, Google/Anthropic). In general, these types of arrangements are therefore pro-competitive and allow businesses to improve their offerings to consumers.[72] From a competition law perspective, specific types of agreements in the AI industry likely fall under Article 101(3) TFEU, which would mean that they are agreements exempted from prohibition under Article 101(1) TFEU.[73] The costs of building the necessary computing in-house capacity would otherwise be extremely high for these smaller firms.

Nevertheless, and as the CMA additionally points out, such agreements could also warrant close scrutiny. There are risks that incumbent firms use those agreements to maintain their position by pre-emptively suppressing potential competitive effects, as is discussed in more detail in Section 3. AI partnerships could increase the risks of foreclosure given incumbents' presence at different points in the value chain,[74] or collusion, particularly when partners share sensitive information about future innovations and business strategies. Such information sharing could inadvertently pave the way for behaviours that dampen competition by limiting innovation or product differentiation, as cautioned by the ADLC.[75] We will turn to the options available to competition authorities to address AI partnership concerns in Section 4.

2.4.4 A global race with regional implications

The development of AI is intrinsically linked to a broader, global race for political leadership through AI.[76] AI has emerged during a time of increasing geopolitical tensions.[77] Governments are competing both in advancing the efficiency of AI models and in the strategic integration of AI into policymaking processes and defence capabilities. [78] This race aims to strengthen not only economic prowess, but also military strength and state power. For example, AI is being actively employed in the war between Ukraine and Russia.[79]

In this context, the 'the AI "arms race" is [...] is layered atop, rather than distinct from, existing [global] power structures and mechanisms.' [80] The traditional superpowers are the ones

[72] OECD (2024), 'Artificial intelligence, data and competition', OECD Artificial Intelligence Papers, No. 18, OECD Publishing, Paris.
[73] Kowalski, K., Volpin, C. and Zombori, Z. (2024), 'Competition in Generative AI and Virtual Worlds', *Competition Policy Brief*, 3.
[74] Carugati, C. and Kar N. (2024), 'Assessing the competitive dynamics of AI partnerships', November, p. 12.
[75] The risk of collusion topic is out of the scope of this briefing paper. However, the ADLC raised it in its study of the AI sector. See ADLC (2024), 'Generative artificial intelligence: the ADLC issues its opinion on the competitive functioning of the sector', 28 June.
[76] Csernatoni, R. (2024), 'Charting the Geopolitics and European Governance of Artificial Intelligence', *Carnegie Europe*, March, p. 7.
[77] Csernatoni, R. (2024), 'Charting the Geopolitics and European Governance of Artificial Intelligence', *Carnegie Europe*, March, p. 3.
[78] For example, Tlis, F. (2025), 'China uses DeepSeek, other AI models, for surveillance and information attacks on US', Polygraph, 4 March.
[79] Csernatoni, R. (2024), 'Charting the Geopolitics and European Governance of Artificial Intelligence', *Carnegie Europe*, March, pp. 5–6.
[80] Parakilas, J. (2025), 'For Geopolitics, What AI Can't Do Will Be as Important as What It Can', *RAND*, 3 April.

primarily competing with each other in the development of AI to gain global leadership.[81] They employ regulations and trade policy as well as large-scale investments for this purpose:

- The US is currently at the forefront of AI development, attributable to a large number of US-based key players in AI (e.g. Anthropic and OpenAI).[82] It has introduced comparatively little regulation for these companies,[83] and relies on large-scale private and public investments worth trillions of dollars to drive innovation. For instance, Nvidia recently pledged $500bn to producing American AI servers, while the US government is investing $39bn in semiconductor facilities through the 'CHIPS for America Fund'.[84] At the same time, the US has prohibited the sale of chips above a certain capability threshold to China, and banned the sale of specific chip manufacturing equipment.[85]
- China is emerging as a key player in AI model optimisation (e.g. DeepSeek).[86] Faced with trade restrictions by the EU and US, it appears to have adopted a 'fast follower' approach of strategic state-led investment that largely focuses on improving existing AI models.[87] In total, China has committed $1.4trn over the next 15 years to become the world leader in AI technology.[88]
- The EU is trailing behind in AI development as of now.[89] Currently it only hosts a few significant AI developing companies,[90] and has focused on positioning itself as an 'ethical leader' in AI by introducing stringent regulation.[91] The EU aims to close the gap with the US and China in AI development, notably through a €200bn investment under the InvestAI initiative.[92]
- Beyond these superpowers, there are also some notable other countries that have started to invest substantial funds in AI. For example, Saudi Arabia created a $40bn fund to invest in AI technology, while the UAE aims 'to build the largest

[81] Parakilas, J. (2025), 'For Geopolitics, What AI Can't Do Will Be as Important as What It Can', *RAND*, 3 April.
[82] Shrivastava, R. (2025), 'AI 50', Forbes, 10 April, accessed 4 June 2025.
[83] Schweizerische Eidgenossenschaft (2024), 'Analysis of artificial intelligence regulation in countries and regions around the world', baseline analysis, 16 December, p. 8.
[84] TrendForce (2025), '[News] NVIDIA Plans $500 billion AI Investment in U.S., with Taiwan's TSMC and Foxconn Leading the Charge', 15 April; Holland, S. (2025), 'Trump announces private-sector $500 billion investment in AI infrastructure', Reuters, 22 January; Lu, M. (2025), 'Visualising Global Investment by Country', Visual Capitalist, 21 April; Kurilla M. (2024) 'What is the CHIPS Act?', Council of Foreign Relations, 29 April (accessed 4 June 2025, respectively).
[85] Bureau of Industry and Security (2022), 'Commerce Implements New Export Controls on Advanced Computing and Semiconductor Manufacturing Items to the People's Republic of China (PRC)', 7 October.
[86] Yang, Z. (2025), 'How Chinese AI StartUp DeepSeek Moade a Model that Rivals OpenAI', Wired, 25 January, accessed 4 June 2025.
[87] Podda, L. (2025), 'China's Drive to Dominate the AI Race', The Atlas Institute for International Affairs, 14 April, accessed 4 June 2025.
[88] Bicker, L. (2025), 'From chatbots to intelligent toys: How AI is booming in China', BBC, 10 March, accessed 4 June 2025.
[89] Csernatoni, R. (2024), 'Charting the Geopolitics and European Governance of Artificial Intelligence', *Carnegie Europe*, March, p. 7.
[90] Shrivastava, R. (2025), 'AI 50', Forbes, 10 April, accessed 4 June 2025.
[91] European Commission (2025), 'AI Act', accessed 3 June 2025.
[92] European Commission (2025), 'EU launches InvestAI initiative to mobilise €200 billion of investment in artificial intelligence', 11 February, accessed 4 June 2025.

artificial intelligence campus outside the US' for better access to advanced AI chips.[93]

While the AI arms race is set to shape AI development, we do not explore its long-term impacts on competition as part of this briefing paper. It is clear, however, that companies are likely to be attracted by large-scale investments in AI made at the national level. Similarly, AI developing firms will generally consider countries' regulation before choosing to offer their product in a particular jurisdiction. For example, the EU AI Act is said to be the reason why Meta's LlaMA models are not available in the EU.[94] These factors are likely to shape the regional development of AI to the advantage, at least in the short term, of countries with less regulation and high financial incentives.

[93] Farrell, M. and Copeland, R. (2024), 'Saudi Arabia Plans $40 Billion Push Into Artificial Intelligence', The New York Times, 19 March, accessed 4 June 2025; The Guardian (2025), 'Trump agrees deal for UAE to build largest AI campus outside the US', 16 May.
[94] the choice (2024), 'How Europe's AI Act could affect innovation and competitiveness', 4 July, ESCP Business School.

3 Is there a need to intervene in AI markets to protect competition?

Amid the calls for intervention in AI-related markets to protect competition, it is worth stopping to consider from a principled basis: 'When is intervention needed in a market?'. The first and most basic grounding for any intervention should be evidence of a market failure, with a demonstrable potential for harm to consumer welfare relative to the expected outcome without intervention.[95]

In the remainder of this briefing paper we focus on assessing whether the dynamics observed in the AI market are evidence of competition working well or of the presence of market failures. Our focus will be on market power failures (i.e. when monopolies or oligopolies restrict output, anti-competitively affect prices, or limit competition) which could result in reduced innovation in the creation of inputs needed for AI model development, higher prices for access to AI inputs, and models or restricted choice of products and services using generative AI.[96] This review draws on the assessments undertaken by competition authorities in the UK, France, Portugal, and Hungary that have conducted market studies to assess competition in generative AI markets. They have identified both current and potential issues in generative AI sectors, ranging from market dominance and unfair practices to consumer protection concerns. While some findings highlight existing problems, others focus on potential risks that may require proactive measures to ensure effective competition in the future.[97]

This section will present an assessment of:

i. the economic characteristics of markets across the AI value chain (network and data feedback effects, multihoming, economies of scope and scale) and their potential to lead to a market tipping (section 3.1);
ii. firm behaviours that could give rise to competition concerns (lock-in and foreclosure in cloud computing; refusal of or discriminatory access to proprietary data; self-preferencing, tying and bundling; and refusal to give access in downstream markets) (section 3.2).

3.1 AI market features—potential for tipping

The impact on competition from a nascent technology such as AI is difficult to predict in the long term. On the one hand, the emergence of new products and services as described in

[95] Alexiadis, P. and De Streel, A. (2020), 'Designing an EU intervention standard for digital platforms', working paper.
[96] Other market failures not covered include the presence of externalities leading to over- or under-provision, under-provision of public goods, and information asymmetries that distort choices and reduce market efficiency.
[97] CMA (2024), 'AI Foundation Models - Update Paper', 11 April; ADLC (2024), 'Generative artificial intelligence: the ADLC issues its opinion on the competitive functioning of the sector', 28 June; Portuguese Autoridade da Concorrência (2023), 'Competition and Generative Artificial Intelligence - Issues Paper', November; Hungarian competition authority (2024), 'Examining the impact of artificial intelligence on market competition and consumers', October.

Section 2 gave rise to vibrant competition that has emerged in a rather short period of time at all levels of the value chain.[98]

At the same time, there is a growing concern that the rapid development and deployment of generative AI could lead to anti-competitive market structures across the AI value chain. Factors such as network effects and cumulative data advantages, economies of scale or scope, and single-homing practices have historically been identified as contributing to the market tipping in favour of certain firms, particularly in technology-driven markets.[99] The AI value chain exhibits several of these factors. However, the extent to which they may lead to increased concentration and tipping depends on: (i) the strength of each effect, and (ii) their cumulative effect within a layer of the value chain. Table 3.1 below provides an overview of the strength of different market features across the AI value chain; the discussion that follows highlights the main insights derived from the overview.

Table 3.1 Current strength of market features across the AI value chain

	Inputs			Modelling		Deployment
Market feature	**Computing**	**Data**	**Talent**	**FM creation**	**Fine-tuning**	**Deployment**
Network effects*	Medium	Low	Low	Medium	Low	Medium
Data feedback effects	Low	Medium	n/a	Low	Medium	Medium
Single-homing	Medium	Low	Medium	Medium	Low	Low
Economies of scale	High	Medium	Low	Medium	Low	Medium
Economies of scope	Medium	Medium	Low	Medium	Low	High

Note: *Data feedback loops are a type of network effects, however, in this table they are treated as a separate market feature.
Source: Oxera analysis based on Martens, B. (2024), 'Why artificial intelligence is creating fundamental challenges for competition policy', Bruegel, 18 July; Hagiu, A. and Wright, J. (2025) 'Artificial intelligence and competition policy', *International Journal of Industrial Organization*, 103134; DCCA (2025), 'Market tipping: Guidance for competition assessments', May; Korinek, A. and Vipra, J. (2025), 'Concentrating intelligence: scaling and market structure in artificial intelligence' *Economic Policy*, **40**:121, pp. 225–256.

First, network effects that do not involve data feedback loops, at any layer of the AI value chain, are not as strong as in other digital markets (e.g. operating systems or marketplaces) since participants' interactions are less likely to lead to improvements to a product or service. Unlike marketplaces, where adding users is nearly cost-free, scaling AI requires incremental computational resources and infrastructure, which limits the direct benefits of additional users. Network effects are also stronger when there are direct interactions between users (e.g.

[98] Schrepel, T. and Pentland, A.S. (2024), 'Competition between AI foundation models: dynamics and policy recommendations', Industrial and Corporate Change, **00**, pp. 1–19, p. 1.
[99] Bedre-Defolie Ö. and Nitsche, R. (2020), 'When Do Markets Tip? An Overview and Some Insights for Policy', *Journal of European Competition Law & Practice*, **11**:10, December, pp. 610–622.

social networks) or contribution to a shared ecosystem (e.g. apps on a specific operating system). AI models, however, process data rather than facilitate direct interactions (except in the deployment layer), and this can weaken the traditional feedback loops that reinforce network effects in the inputs and modelling layer.[100] As network effects are considered one of the prime ingredients for market tipping, a priori, the AI value chain seems to be less predisposed to this outcome.

Second, the compute layer of the value chain might be prone to tipping both in the provision of AI chips and the provision of cloud computing infrastructure due to a specific combination of factors including:

- high economies of scale due to capital-intensive production of advanced AI chips, and the high cost of maintenance and development of AI cloud infrastructure to provide sufficient compute power to meet the need for more complex AI model development;[101]
- a medium level of network effects (direct and indirect) as a small number of firms in the input layer benefit from improvements to their products due to a larger user base (e.g. Nvidia's improvements to its CUDA software due to a large user base);[102] and
- a medium level of economies of scope since providers at this layer have a cost advantage from being able to produce products not only for AI and deployment but also related activities (e.g. chips for gaming and collocated general cloud services).[103]

The market position enjoyed by Nvidia[104] and the importance of hyperscalers in the compute layer could, on the one hand, be read as evidence that the market has already tipped. On the other hand, there is an increase in the number of market entrants in AI chip production (as discussed in section 2.3.1) and there are several mitigating factors with regard to the advantage enjoyed by compute providers (as discussed in depth in section 3.2.1). Overall, while this is an area that will likely remain of interest to competition authorities, the long-term potential for a winner-take-all outcome remains uncertain.

Third, data feedback effects could manifest across the entire value chain. However, they are likely to be more significant when they involve proprietary, used-supplied and user-created data used to improve the products and services through continuous learning feedback loops, which takes place at the fine-tuning and deployment layer. Specifically:

[100] OECD (2024), 'Artificial intelligence, data and competition', *OECD Artificial Intelligence Papers*, **18**, pp. 29–30.
[101] CMA (2023), 'AI Foundation Models – Initial Report', 18 September, pp. 33–34.
[102] Hagiu, A. and Wright, J. (2025) 'Artificial intelligence and competition policy', *International Journal of Industrial Organization*, 103134, p. 2.
[103] NVIDIA (2024), 'NVIDIA Brings Generative AI to Millions, With Tensor Core GPUs, LLMs, Tools for RTX PCs and Workstations', 8 January, accessed 29 May 2025.
[104] Hagiu, A. and Wright, J. (2025) 'Artificial intelligence and competition policy', *International Journal of Industrial Organization*, 103134

- at the development layer, the use of publicly available data and limited feedback from the use of the AI model downstream means that there are limited data feedback effects being generated.[105]
- at the fine-tuning layer, while user specific data can play an important role, the risk of markets tipping due to data feedback loops is not considered to be high given that no firm has exclusivity over a category of data.[106]
- at the deployment layer, there is currently a multitude of AI options available to choose from.[107] However, given that vertically integrated firms have an advantage due to downstream pre-existing routes to users—which can offer both historical and continuous data, as well as strong feedback signals—the potential for market tipping at the level of specific services remains a possibility.[108] This risk is enhanced by the advantage offered by economies of scope enjoyed by large tech firms across their entire suite of products and services.

Fourth, currently multi-homing and switching is possible at least to some extent across the value chain. In the input layer, customers are actively trying to diversify the providers and partners they work with (see sections 2.3.1 and 2.4.3). Similarly, users (both individuals and businesses) have access to multiple FMs and fine-tuned models which can be used in parallel or at times together (e.g. Perplexity uses FMs provided by OpenAI and Anthropic).[109] However this might be subject to change in the future. As AI functionality becomes more advanced and embedded in the workflows, products and services of their customers, multi-homing might become more costly or inconvenient, thus raising switching costs.[110] This transition from low-friction experimentation to high-friction integration is a key inflection point that could shift deployment from a contestable space toward one more vulnerable to tipping point; however, it is too soon to know which direction it will take.

Fifth, building FMs involves significant fixed costs and exhibits clear economies of scale,[111] however the current level of competition among many players means tipping is unlikely in the short term. Larger firms can amortise their costs across many applications and customers. Data feedback effects arise as firms with more compute, data, and talent build better models, which, in turn, attract more customers, talent, and usage data. This cycle increases the likelihood of concentration at this layer, where scale is a major competitive advantage. However, as discussed in section 2.3.2, the entry of new players motivated by the large scale of future potential gains has given rise to a dynamic and competitive situation in the FM layer with no current signs of slowing down.

[105] Hagiu, A. and Wright, J. (2025) 'Artificial intelligence and competition policy', *International Journal of Industrial Organization*, 103134, pp. 5–6.
[106] Hagiu, A. and Wright, J. (2025) 'Artificial intelligence and competition policy', *International Journal of Industrial Organization*, 103134.
[107] Hagiu, A. and Wright, J. (2025) 'Artificial intelligence and competition policy', *International Journal of Industrial Organization*, 103134.
[108] CMA (2024), 'AI Foundation Models - Update Paper', 11 April.
[109] Hagiu, A. and Wright, J. (2025) 'Artificial intelligence and competition policy', *International Journal of Industrial Organization*, 103134, p. 5.
[110] DCCA (2025), 'Market tipping: Guidance for competition assessments', May, p. 74.
[111] Hagiu, A. and Wright, J. (2025) 'Artificial intelligence and competition policy', *International Journal of Industrial Organization*, 103134.

Overall, despite the presence of certain economies of scope or scale, as well as network and feedback effects, it is currently uncertain whether the market is likely to tip toward specific firms.. As descried above, the areas with the highest potential for tipping are in the compute layer and certain markets in the deployment layer, where large, vertically integrated incumbents may be able to benefit from the market features that give them a competitive advantage.

In the next section, we consider what types of behaviour might be used by incumbent tech firms in AI markets that could give rise to competition concerns and the potential mitigation factors that might prevent this.

3.2 Main competition concerns related to AI

AI is still a nascent technology with no established theories of harm through decisions by competition authorities or judgments by courts.[112] As of now, the potential concerns that have been identified by competition authorities fit well under the traditional theories of harm and it is unclear whether novel issues will emerge in the future.[113]

In particular, competition authorities are primarily concerned that large incumbent digital tech firms could use their current market position and financial and technological capabilities to establish significant and entrenched market power across the AI value chain.[114] For instance, the ADLC describes that:

'[…] **major digital companies**' access to key inputs and the advantages linked to their vertical and conglomerate integration **create the conditions for strong concentration, to their benefit'**[115] [emphasis added].

Competition authorities therefore see a risk that the generative AI sector may be shaped by digital companies to only have one or a few dominant suppliers, similar to the outcome observed in earlier digital markets. Indeed, as explained above in section 2.3.1, large digital companies already hold a significant position in the input layer of the AI value chain (for example, AI chips and cloud computing).[116] In the modelling layer, as discussed in section 2.3.2, the level of competition is currently high. However, there is a risk that large digital firms may leverage their strong position in either the input or deployment layers to strengthen their position in the modelling layer over the medium to long term.[117] This concern is highlighted by the large number of partnerships and investments by incumbent digital firms taking place

[112] OECD (2024), 'Artificial intelligence, data and competition', background note, 6 May, para. 91.
[113] OECD (2024), 'Artificial intelligence, data and competition', background note, 6 May, para. 90.
[114] CMA (2024), 'AI Foundation Models', technical update report, 16 April, para. 5.3; Competition Bureau Canada (2024), 'Artificial intelligence and competition – Discussion Paper', 20 March, pp. 17–18; AdC (2023), 'Competition and Generative Artificial Intelligence', Issues Paper, November, p. 27.
[115] ADLC (2024), 'opinion 24-A-05 on the competitive functioning of the generative artificial intelligence sector', 28 June, p. 7.
[116] For instance, competition authorities have found that AI firms hold a significant position in cloud computing, see: Longo, R. and Rocha, M. (2024), 'Generate AI: The new digital frontier for competition', *Concurrences no.2-2024*, p. 24; CMA (2025), 'Cloud services market investigation', summary of provisional decision, para. 6.
[117] ADLC (2024), 'opinion 24-A-05 on the competitive functioning of the generative artificial intelligence sector', 28 June, pp. 7 and 9.

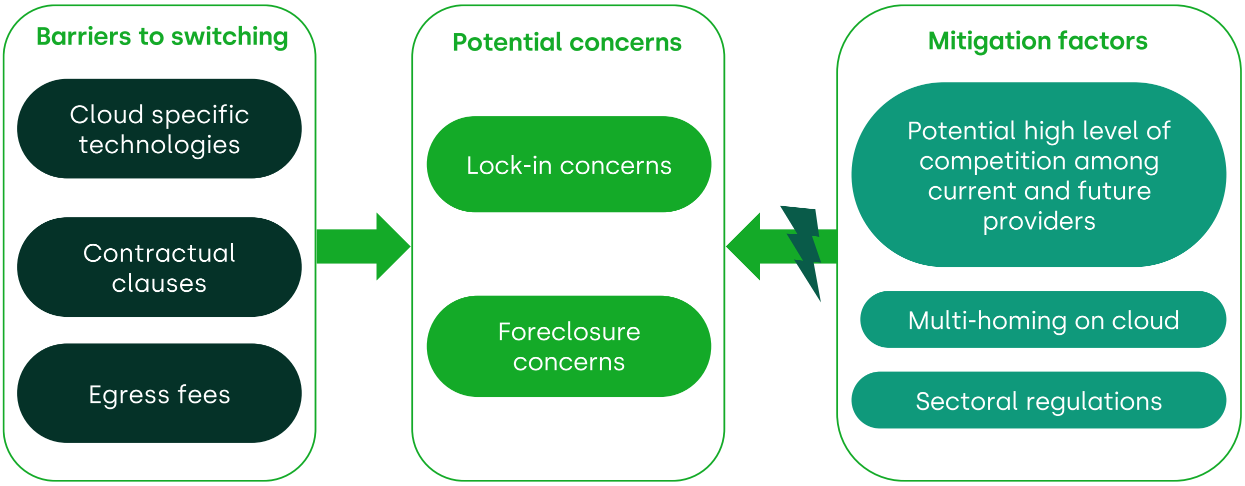
across the AI value chain (see section 2.4.3) that raise several other concerns related to vertical integration.[118]

In this section, we focus on the potential threat to competition posed by large incumbent digital companies in cloud computing, proprietary data and from downstream integration.[119]

3.2.1 Risks in cloud computing

A first layer where competition concerns could arise is in cloud computing (see section 3.1). This layer of the value chain is understood to present technical and financial barriers to switching that may restrict AI developer companies' ability to migrate providers or to multi-home. This has led to concerns that AI developing firms or users may be **locked-in** to using services from certain cloud providers; and that large digital firms, who also develop their own models, or have in place partnerships with specific AI developing firms, may **foreclose access to their cloud computing services** to competitors. These concerns may be somewhat mitigated due to current and future competition among the hyperscalers themselves, multi-homing practices of generative AI developing firms, and sectoral regulations that can be used to facilitate switching between providers. We present each of these points in Figure 3.1 and discuss them in turn below.

Figure 3.1 Competition concerns in cloud computing



Source: Oxera.

[118] ADLC (2024), 'opinion 24-A-05 on the competitive functioning of the generative artificial intelligence sector', 28 June, p. 9.
[119] Other potential theories of harm that are not covered by this briefing paper include: anti-competitive agreements within AI partnerships (e.g. information exchange, algorithmic collusion, cartel-like agreements), anti-competitive labour agreements, and exploitative abuse in the AI chip provision markets.

The risks

Firms often contract cloud computing services to gain access to sufficient computing power to develop generative AI models.[120] As described above in section 2.3.1, the cloud sector is highly concentrated with the presence of three hyperscalers (Amazon, Microsoft and Alphabet).[121] For generative AI providers, it is currently understood to be challenging (but not impossible) to switch CSPs due to technological and financial barriers.

CSPs tend to offer cloud specific technologies that might make migrations more difficult, time-consuming and expensive:

- CSPs may offer proprietary technology for creating or fine-tuning AI models. For instance, AWS offers 'AWS Health' and 'HealthLake' which help AI models transform medical and insurance data, while GCP facilitates medical diagnoses for AI models through 'Cloud Healthcare API' or 'Medical Imaging Suite'.[122] Such tools generally benefit consumers by providing specialised solutions. However, they also make models difficult to migrate since they would need to be recreated from scratch when using a different provider.[123]
- CSPs may use unique AI chip types, as highlighted in section 2.3.1. These can be difficult to switch away from, as some AI models rely on specific AI chip features and programming languages (such as Nvidia's CUDA) and would need a substantial software re-write to migrate providers.[124] For example, AWS's Neuron Software Development Kit is currently only compatible with AWS's Inferentia and Trainium chips.[125]

In addition, CSPs may also adopt policies that could act as a further barrier to switching for AI developing firms if they enter into partnership contracts that disincentivise switching. As set out in section 2.4.3, many model developing firms are reliant on partnerships with CSPs to access the necessary computing power to develop their models.[126] These partnerships could include clauses that explicitly forbid AI companies to multi-home, or they might provide implicit incentives to not switch between CSPs by requiring spending commitments on the partner's cloud services. These spend commitments are very common and could limit the financial capabilities of AI developing companies to contract other CSPs.[127]

[120] Longo, R. and Rocha, M. (2024), 'Generate AI: The new digital frontier for competition', *Concurrences no.2-2024*, p. 24.
[121] As mentioned in section 2.3.1, the CMA recently found that AWS and Microsoft together hold a market share in the UK of between 70% and 90%, see: CMA (2025), 'Cloud services market investigation', summary of provisional decision, para. 6.
[122] Van der Vlist, F., Helmond, A. and Ferrari, F. (2024), 'Big AI: Cloud infrastructure dependence and the industrialisation of artificial intelligence', *Big Data & Society*, **11**:1, p. 10.
[123] ADLC (2024), 'opinion 24-A-05 on the competitive functioning of the generative artificial intelligence sector', 28 June, paras 246–252.
[124] Federal Trade Commission (2025), 'Partnerships Between Cloud Service Providers and AI Developers', staff report, January, p. 34.
[125] Federal Trade Commission (2025), 'Partnerships Between Cloud Service Providers and AI Developers', staff report, January , p. 34.
[126] OECD (2024), 'Artificial intelligence, data and competition', OECD Artificial Intelligence Papers, No. 18, OECD Publishing, Paris.
[127] Federal Trade Commission (2025), 'Partnerships Between Cloud Service Providers and AI Developers', staff report, January , pp. 31 and 32.

CSPs may also adopt a pricing structure that, on the one hand, benefits users through lower initial prices and lower costs due to discounts, but on the other, further increases the costs of switching cloud providers. Specifically:

- CSPs may adopt egress fees which 'charge per outgoing bandwidth usage',[128] in which case the fees increase proportionally to the volume of data transferred to other CSPs.[129] These types of fees are often justified as there are costs involved in moving data from a cloud.[130] While egress fees can be used by all types of CSPs, it is mainly hyperscalers that have introduced them in their pricing structure.[131] For the AI developing firms buying compute from a CSP that charges egress fees, the fees may pose a significant barrier to migrating providers.
- Hyperscalers have also made it particularly attractive for companies to take up their services. Specifically, they tend to offer large discounts in the form of cloud credits. These cloud credits are common among all CSPs and typically grant free or cheaper access to the CSP's services for a defined period. For instance, customers that want to test new functionalities currently have the following cloud credit offers available: $1,000 without time limited for AWS and $300 over 60 days for GCP, compared to 20 EUR and 30 EUR for the services offered by the smaller CSPs Clever Cloud and Qarnot Computing, respectively.[132] Startups additionally have access to $200,000 in credits over two years from GCP, $150,000 from Microsoft Azure, and $100,000 from AWS while most smaller CSPs tend to offer significantly less (or nothing at all).[133]

These barriers together risk leading to a **technical and financial lock-in** of companies once they use a cloud provider's infrastructure. This concern might be especially present in the case of hyperscalers which offer proprietary solutions for creating or fine-tuning AI models, charge egress fees, and have invested extensively in exclusive or strategic partnerships that limit AI developing firms' ability to switch providers.[134]

Another concern in relation to hyperscalers is that they may engage in **input foreclosure** practices to gain a stronger position at the modelling level of the AI supply chain. The main providers of cloud computing services all develop their own generative AI models and thereby help to create incremental value by increasing the supply of providers to choose from (see section 2.3.1). At the same time, they also have partnerships with other generative AI

[128] ADLC (2023), 'opinion 23-A-08 on competition in the cloud sector', 29 June, p. 127.
[129] ADLC (2023), 'opinion 23-A-08 on competition in the cloud sector', 29 June, p. 127.
[130] Biglaiser, G., Crémer, J. and Mantovani, A. (2024), 'The Economics of the Cloud', Toulouse School of Economics Working Papers, 1520, p. 25.
[131] ADLC (2023), 'opinion 23-A-08 on competition in the cloud sector', 29 June, p. 127.
[132] ADLC (2023), 'opinion 23-A-08 on competition in the cloud sector', 29 June, p. 59.
[133] An exception to this are Clever Cloud and OVH, which both offer 100,000 EUR in progressive rebates or credits, see: ADLC (2023), 'opinion 23-A-08 on competition in the cloud sector', 29 June, p. 59.
[134] ADLC (2024), 'opinion 24-A-05 on the competitive functioning of the generative artificial intelligence sector', 28 June, paras 246–252; ADLC (2023), 'opinion 23-A-08 on competition in the cloud sector', 28 June, p. 127.

developing companies. As such, they may have an incentive to foreclose the developers of competing models from accessing their cloud computing services.[135]

Indeed, model developing firms still seem to be dependent on partnerships with hyperscalers to access sufficient computing power. As of 2023, smaller CSPs reported that they were struggling to meet the demands made by AI developers, and it is unclear how far this situation has changed since.[136] This could have a greater effect on model developing firms that frequently partner with smaller or specialised CSPs, such as CoreWeave.[137] However, hyperscalers continue to provide greater storage and computing capacities than these companies. As a result, during the AI chip shortages in 2023, AI developing firms reported not being able to get access to chips 'unless you [AI developing firm] have some existing contract with [hyperscalers] or you're prepaying for it.'[138]

Potential mitigating factors

In terms of mitigating factors, the competition between the hyperscalers may prevent or mitigate foreclosure practices. Currently, there are three hyperscalers present in cloud computing.[139] These hyperscalers compete intensely with each other, with a study showing that 'consumers have been having access to more innovative products with increasing quality at lower prices'.[140] In principle, generative AI developers should therefore be able to choose several other providers if one hyperscaler decides to foreclose its services.

However, as discussed above, AI developing companies may be restricted in their ability to react to foreclosure practices after having entered a cloud partnership due to the switching barriers mentioned above. To mitigate this concern, model developing firms have started partnering widely with multiple infrastructure providers.[141] Since partnerships do not tend to be exclusive, there is usually at least the option for AI developing firms to buy cloud computing services from several suppliers. For instance, OpenAI has recently put in place partnerships with SoftBank and Oracle among others to mitigate sole reliance on Microsoft.[142] Similarly, Anthropic has a partnership with both AWS and Google Cloud.[143]

In addition, migration of services across clouds may become easier in the future. There are some platforms being developed that facilitate data transfers and the management of source codes for AI developers (e.g. Kubernetes or OpenStack). These are currently mainly focused

[135] Bostoen, F. and van der Veer, A. (2024), 'Regulatory Competition in Generative AI: A Matter of Trajectory, Timing and Tools', *Concurrences no.2-2024*, p. 31.
[136] Federal Trade Commission (2025), 'Partnerships Between Cloud Service Providers and AI Developers', staff report, January, p. 31.
[137] Businesswire (2023), 'Inflection AI Announces $1.3 Billion of Funding Led by Current Investors, Microsoft, and NVIDIA', accessed 15 May 2025; MSV, J. (2024), 'Vultr and Runi:ai Partner to Enhance AI Infrastructure', Forbes, 31 July, accessed 23 May 2025.
[138] Holmes, A. and Gardizy, A. (2023), 'AI Developers Stymied by Server Shortage at AWS, Microsoft, Google', The Information, 7 April, accessed 23 May 2025.
[139] ADLC (2023), 'opinion 23-A-08 on competition in the cloud sector', 29 June, p. 4.
[140] Parisi, R. (2024) 'The Cloud Services Markets' Competitive Landscape: A contribution to the Competition and Markets Authority', *Competition & Innovation Lab*, p. 3.
[141] Bostoen, F. and van der Veer, A. (2024), 'Regulatory Competition in Generative AI: A Matter of Trajectory, Timing and Tools', *Concurrences no.2-2024*, p. 31.
[142] Wiggers, K. (2025), 'Microsoft is no longer OpenAI's exclusive cloud provider', TechCrunch, accessed 28 May 2025.
[143] Anthropic (2024), 'Anthropic, AWS, and Accenture team up to build trusted solutions for enterprises', 20 March; Anthropic (2023), 'Anthropic Partners with Google Cloud', 3 February (accessed 28 May 2025, respectively)

on increasing interoperability between different cloud systems and cannot currently be used to migrate AI models that use proprietary tools offered by CSPs.[144] However, the latter cases might be limited due to the multi-homing behaviour of AI developing firms and the potential for the growing adoption of cross-cloud solutions to incentivise the development of technologies that support broader AI migration capabilities in the future.

Also, predicted market developments are likely to further increase competition in cloud computing, and thereby impede hyperscalers' ability to foreclose effectively. Specifically, the AI industry is set to grow strongly in the coming years. For instance, Bloomberg has estimated that the generative AI market will be worth $1.3trn by 2032.[145] This means that it is likely to be profitable for smaller CSPs to invest in sufficient capabilities to provide their services to AI companies at a large scale. As a result, the range of competitive cloud service offerings may grow further in the coming years.

Finally, there exist sectoral regulations in several jurisdictions that can be used to facilitate switching between cloud providers and make it easier to multi-home. In the EU, the Digital Markets Act (DMA) prohibits the use of technical or other means to restrict switching for so-called ´gatekeepers´ (i.e. large companies with an entrenched position).[146] While there has not yet been a gatekeeper designated for cloud services, it is possible that Amazon, Microsoft and Google might be designated in the future if their cloud computing services meet the designation criteria.[147] In the UK, the CMA has similar powers under the Digital Markets, Competition and Consumers Act (DMCCA) and can decide to intervene in the provision of cloud computing (including services specialised for AI).[148] Moreover, despite the benefits of having non-linear pricing, under the EU Data Act, the EU has imposed a price-cap on egress fees based on incurred costs until 12 January 2027, after which companies are prohibited from charging them entirely.[149] In all other jurisdictions, companies may continue to charge egress fees.[150, 151]

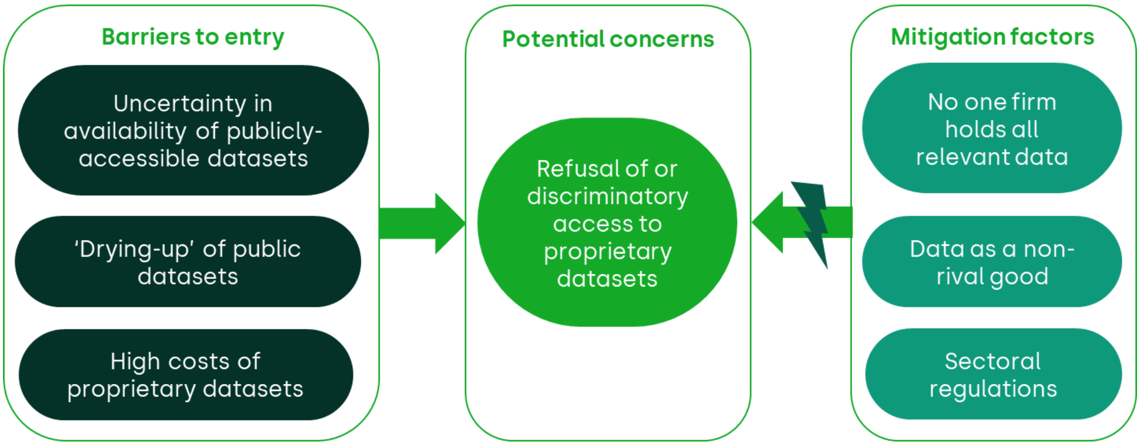
3.2.2 Risks relating to proprietary data

A second input in AI development where competition concerns are particularly high is data (see Figure 3.2). Access to proprietary datasets is likely to become a barrier to develop generative AI models since these datasets are increasingly needed to compete effectively. Large digital companies are enjoying preferential access to proprietary datasets. A frequently highlighted concern is therefore that large digital firms may eliminate competition by restricting, or providing discriminatory access to, a large number of proprietary datasets in the future ('**refusal of access'** or '**discriminatory access**'). Large digital companies' ability to

[144] ADLC (2023), 'opinion 23-A-08 on competition in the cloud sector', 29 June, p. 139.
[145] Bloomberg (2023), 'Generative AI to Become a $1.3 Trillion Market by 2032, Research Finds', 1 June, accessed 28 May 2025.
[146] Autoriteit Consument & Market (2023), 'Market Study Cloud services', 5 September, p. 70.
[147] Autoriteit Consument & Market (2023), 'Market Study Cloud services', 5 September, p. 69.
[148] CMA (2024), 'Digital Markets Guidance', guidance, 19 December, para. 4.23.
[149] von Baum, F. and Gärtner, M.R. (2025), 'Switching and porting rules under the EU Data Act', Pinsent Masons, 20 March, accessed 4 June 2025.
[150] Biglaiser, G., Crémer, J. and Mantovani, A. (2024), 'The Economics of the Cloud', Toulouse School of Economics Working Papers, 1520, p. 23.
[151] ADLC (2023), 'opinion 23-A-08 on competition in the cloud sector', 29 June, p. 127.

effectively refuse access to specialised datasets will likely be somewhat restricted, though, by the availability of data and existing sectoral regulations.

Figure 3.2 Competition concerns in data



Source: Oxera.

We discuss each of these points in turn below.

The risks

As mentioned in section 2.3.1, generative AI models need to be trained on very large, high-quality datasets to create high-quality outputs.[152] For this purpose, model developers generally rely on a mix of publicly available datasets and proprietary datasets including, for example, licensed or in-house data.[153]

In relation to publicly available data which contain copyrighted materials, generative AI developers are facing uncertainty on whether future access to them may require additional licence fees (see section 2.3.1) and hence increase the cost of access to data.[154, 155] Unless it is recognised in law that using copyrighted content for training AI is exempt from copyright rules—e.g. in the EU, there are exceptions to copyrights for AI developing companies unless copyright holders explicitly opted-out[156]—the copyright holders might expect to receive

[152] Kowalski, K., Volpin, C. and Zombori, Z. (2024), 'Competition in Generative AI and Virtual Worlds', *Competition Policy Brief*, 3, p. 4.
[153] Longo, R. and Rocha, M. (2024), 'Generative AI: The new digital frontier for competition', *Concurrences no.2-2024*, p. 25.
[154] OECD (2024), 'Artificial intelligence, data and competition', OECD Artificial Intelligence Papers, 18, p. 31.
[155] Martens, B. (2024), 'Why artificial intelligence is creating fundamental challenges for competition policy?', *Policy Brief*, No. 16/2024, Bruegel, p. 6.
[156] Martens, B. (2024) 'Economic arguments in favour of reducing copyright protection for generative AI inputs and outputs', Working Paper, September, Bruegel, p. 2, accessed 28 May 2025.

licence fees.[157] In fact, AI developing firms have already started to face lawsuits for their use of copyrighted material without prior authorisation, e.g. *Getty Images v Stability AI* in the UK.[158] In the US, there have been multiple lawsuits to clarify whether the use of data without prior permission to train GenAI models falls under the 'fair-use' exception in US copyright, e.g. *The New York Times v OpenAI, Inc. et al.* and *Getty Images (US), Inc. v Stability AI, Inc.*[159]

Depending on the outcomes of legal cases across jurisdictions, content that is currently accessible in public datasets may increasingly require licensing fees, i.e. become proprietary.[160] Indeed, this development already seems to be taking place to some degree. An MIT-led research group found that publishers have started implementing protections to prevent their data from being mined for AI training.[161] Specifically, many copyright holders have explicitly introduced opt-outs that make it illegal to mine their data—for example, under the EU AI Act.[162] Others have set up paywalls or attempted to block web crawlers used by AI developers.[163]

These developments will likely increase the importance of proprietary datasets that contain exclusive data in the coming years.[164] Proprietary datasets are now considered 'the real differentiator'[165] in developing high-quality and specialised AI models. Trusted providers of high-quality data, such as publishers and news outlets, may charge substantial fees for the use of their datasets.[166] For instance, Alphabet recently bought access to data from Reddit for one year, reportedly for around EUR 55m.[167]

AI developing firms are therefore likely to face significant barriers in the future from training data costs becoming increasingly expensive. In this context, large incumbent digital companies possess significant efficiency advantages due to their substantial financial resources, and the wealth of exclusive data generated by their services in downstream markets. Specifically, these companies:

- Can afford to enter data licensing agreements with third-party providers, such as Reddit or The New York Times.[168]

[157] Martens, B. (2024), 'Why artificial intelligence is creating fundamental challenges for competition policy?', *Policy Brief*, No. 16/2024, Bruegel, p. 6.
[158] Davies, C.W. and Dennis, G. (2024), 'Getty Images v Stability AI: the implications for UK copyright law and licensing,' Pinsent Masons, 29 April, accessed 28 May 2025.
[159] Longo, R. and Rocha, M. (2024), 'Generative AI: The new digital frontier for competition', *Concurrences no.2-2024*, p. 25.
[160] OECD (2024), 'Artificial intelligence, data and competition', OECD Artificial Intelligence Papers, 18, p. 31.
[161] Roose, K. (2024), 'The Data That Powers A.I. Is Disappearing Fast', *The New York Times*, 19 July.
[162] Martens, B. (2024), 'Why artificial intelligence is creating fundamental challenges for competition policy?', *Policy Brief*, No. 16/2024, Bruegel, p. 6.
[163] Roose, K. (2024), 'The Data That Powers A.I. Is Disappearing Fast', The New York Times, 19 July.
[164] OECD (2024), 'Artificial intelligence, data and competition', OECD Artificial Intelligence Papers, 18, p. 31.
[165] Andy Thurai in Adebayo, K.S. (2025), 'Why Proprietary Data Is The New Gold For AI Companies', Forbes, accessed 2 May 2025.
[166] Kowalski, K., Volpin, C. and Zombori, Z. (2024), 'Competition in Generative AI and Virtual Worlds', *Competition Policy Brief*, 3, p. 4; Longo, R. and Rocha, M. (2024), 'Generative AI: The new digital frontier for competition', *Concurrences no.2-2024*, p. 25.
[167] ADLC (2024), 'opinion 24-A-05 on the competitive functioning of the generative artificial intelligence sector', 28 June, paras 197–199.
[168] Martens, B. (2024), 'Why artificial intelligence is creating fundamental challenges for competition policy?' *Policy Brief*, No. 16/2024, Bruegel, p. 6.

- Generally enjoy greater access to proprietary datasets from the downstream platforms that they operate. For example, in the case of Alphabet, these include Google Chrome, Google Search and YouTube. Of these, YouTube alone allegedly hosts over 10bn videos.[169]
- Have access to the data generated by their own AI models (e.g. Microsoft through their productivity services, see section 3.2.3). This can create a positive feedback loop where the use of the companies' AI models generates new data, which in turn improves the functioning of their models.[170]

These efficiencies enable large digital companies to more easily provide a high-quality service to consumers. However, they have also led to concerns that the large digital companies may effectively eliminate competition in the future by **refusing** or providing **discriminatory access** to large, specialised datasets that they, or their partners, own.[171] Specifically, there is a risk that large digital companies may deny access to data to competitors, or compel their partners not to sell any data to competitors.[172] Refusing to provide access to its data is, in principle, a company's right and may be due to various reasons, including commercial, technical, or privacy policy obstacles.[173] In this case, there is a risk that the data that large digital companies have access to may become essential to compete in the future.[174]

Indeed, even if competition is not eliminated entirely, smaller companies may be excluded from the market in the future if large digital firms decide not to share the proprietary datasets that they have access to.[175] In particular, their model quality may not become sufficiently high if particularly valuable proprietary datasets become unavailable to them.[176] In that case, the level of competition in the modelling layer of the AI value chain may be negatively affected in the long run.

Potential mitigating factors

Currently, it seems impossible for large digital firms to foreclose all competition in AI models by refusing access to the training data they, or their partners, own.[177] This is firstly because no one firm has exclusive access to sufficient data to effectively eliminate competition by denying access to this data. As described above, Alphabet, Microsoft and Meta all have access to their own proprietary datasets and have bought access to licensed datasets.[178] This

[169] ADLC (2024), 'opinion 24-A-05 on the competitive functioning of the generative artificial intelligence sector', 28 June, paras 197–199.
[170] ADLC (2024), 'opinion 24-A-05 on the competitive functioning of the generative artificial intelligence sector', 28 June, para. 198.
[171] ADLC (2024), 'opinion 24-A-05 on the competitive functioning of the generative artificial intelligence sector', 28 June, para. 255.
[172] ADLC (2024), 'opinion 24-A-05 on the competitive functioning of the generative artificial intelligence sector', 28 June, para. 255.
[173] OECD (2023), 'Theories of harm for digital mergers', OECD Competition Policy Roundtable Background Note, p. 23.
[174] OECD (2023), 'Theories of harm for digital mergers', OECD Competition Policy Roundtable Background Note, p. 21.
[175] Bostoen, F. and van der Veer, A. (2024), 'Regulatory competition in generative AI: A matter of trajectory, timing and tools', *Concurrences no.2-2024*, p. 31.
[176] Bostoen, F. and van der Veer, A. (2024), 'Regulatory competition in generative AI: A matter of trajectory, timing and tools', *Concurrences no.2-2024*, p. 31.
[177] Bostoen, F. and van der Veer, A. (2024), 'Regulatory competition in generative AI: A matter of trajectory, timing and tools', *Concurrences no.2-2024*, p. 31.
[178] ADLC (2024), 'opinion 24-A-05 on the competitive functioning of the generative artificial intelligence sector', 28 June, paras 197–199.

gives each of them the opportunity to compete effectively, even if they do not have access to each other's datasets. Secondly, training data is a non-rival good—i.e. a good that many companies can use simultaneously.[179] This means that several large, high-quality datasets suffice to successfully train a very large number of AI models.[180] As such, AI companies will be able to compete as long as a select few companies make their high-quality, proprietary datasets available. These include publishing companies that have access to a wealth of high-quality data, or smaller providers of comparable data to that owned by large digital companies.[181] For instance, collaborative interaction data from Slack could substitute for the same type of data from Microsoft Teams.[182] Thirdly, there are also open-source models available that can be used to train new models, thereby avoiding the need to access certain initial datasets.[183]

Additionally, large digital companies may not have a financial incentive to exclude smaller companies by refusing access to their proprietary datasets. This will depend on whether they expect to profit more from the expanded use of their AI models downstream than from allowing access to their data upstream.[184] In general, making data available is associated with low direct costs.[185] The margin earned on the upstream sale of data may therefore be substantial (depending on the prices charged) and could outweigh expected loss of profits downstream.

More generally, competition authorities would also have the power to intervene in these cases through sectoral digital regulation when the concern is related to a designated firm in scope of specific digital regulation. In the context of the EU, the DMA gives the Commission the ability to prohibit the use of personal data or impose data sharing obligations on gatekeepers for downstream services, i.e. Alphabet, Amazon, Apple, Meta and Microsoft, among others.[186] The Commission has explicitly clarified that the DMA provisions can be applied on any AI system embedded by these gatekeepers in a core platform service (e.g. Gemini embedded in Google Search, or Meta AI employed by Meta in WhatsApp) which could attract obligations for access interoperability to relevant data from the regulated service.[187] Indeed, the Commission has already stated that certain types of personal and business data may not be used by

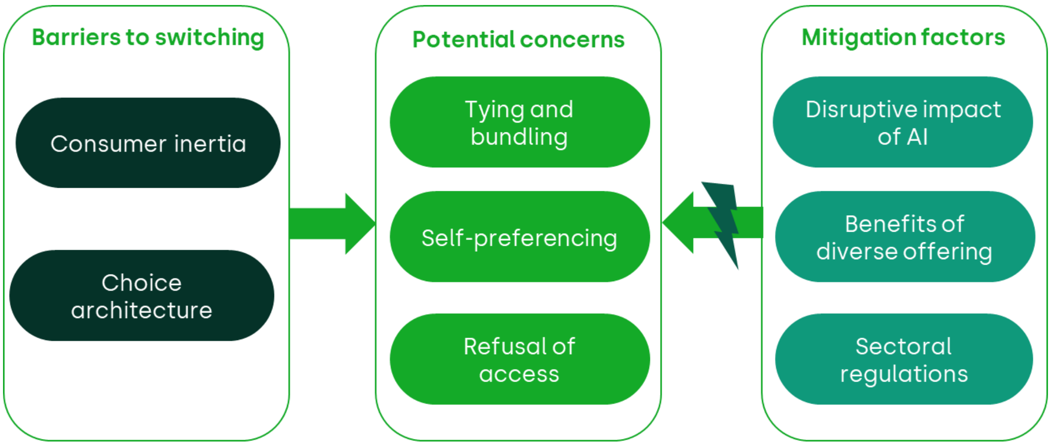
[179] Martens, B. (2024), 'Why artificial intelligence is creating fundamental challenges for competition policy?' *Policy Brief*, No. 16/2024, Bruegel, p. 9.
[180] Longo, R. and Rocha, M. (2024), 'Generative AI: The new digital frontier for competition', *Concurrences no.2-2024*, p. 25.
[181] Martens, B. (2024), 'Why artificial intelligence is creating fundamental challenges for competition policy?' *Policy Brief*, No. 16/2024, Bruegel, p. 6; OECD (2023), 'Theories of harm for digital mergers', OECD Competition Policy Roundtable Background Note, p. 24.
[182] Schelble, B.G., Lancaster, C., Mallick, R., McNeese, N.J., Freeman, G., and Pak, R. (2024), 'A Comparative Evaluation of Ad Hoc Team Performance, Effectiveness, and Interactions in Modern Collaborative Technology', *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 68, p. 1259.
[183] Martens, B. (2024), 'Why artificial intelligence is creating fundamental challenges for competition policy?' *Policy Brief*, No. 16/2024, Bruegel, p. 4.
[184] Guidelines on the assessment of non-horizontal mergers under the Council Regulation on the control of concentrations between undertakings, Official Journal of the European Union, 2008/c 265/07, para. 40.
[185] Houghton, J. (2011), 'Costs and Benefits of Data Provision', Centre for Strategic Economic Studies Victoria University, p. 8.
[186] European Commission (2025), 'The Digital Markets Act: Ensuring fair and open digital markets', accessed 16 May 2025; Langus, G., Maier, N. and Muhamedrahimov, R. (2024), 'Balanced and Transparent Antitrust in the AI Space', *Concurrences no.2-2024*, p. 14.
[187] European Commission (2024), 'High-Level Group for the Digital Markets Act Public Statement on Artificial Intelligence', 22 May; WhatsApp (2025), 'About Meta AI'; Google (2024), 'Generative AI in Search; Let Google do the searching for you', 14 May (accessed 28 May 2025, respectively).

gatekeepers 'in competition with business users'.[188] The CMA also has similar powers under the DMCCA.[189]

3.2.3 Risks from downstream integration

Finally, the downstream integration of large digital companies might give rise to competition risks in specific circumstances (see Figure 3.3). The concern here is that large, vertically integrated firms may be able to distort competition between AI models downstream by leveraging their market position in the downstream markets where they are already significant players. Specifically, they may increase the use of their own generative AI models through **tying and bundling, self-preferencing** or **refusing to grant access** to their products and services downstream for integration with other AI models.[190] Nevertheless, these leveraging risks may be mitigated by the potential disruptive impact of AI in the incumbents' markets, incentives to provide choice to users, or by existing sectoral regulations.

Figure 3.3 Competition concerns from the downstream integration of large digital companies



Source: Oxera.

Each of these points are discussed in more detail below.

The risks

Alphabet, Meta, Amazon and Microsoft are all vertically integrated, and benefit from partnerships across the AI value chain. Upstream, they may provide cloud computing services, access to open-source models, or sell chips or data to other generative AI companies (see

[188] European Commission (2024), 'High-Level Group for the Digital Markets Act Public Statement on Artificial Intelligence', 22 May, accessed 28 May 2025.
[189] CMA (2024), 'Digital Markets Guidance', guidance, 19 December, para. 4.23.
[190] Kowalski, K., Volpin, C. and Zombori, Z. (2024), 'Competition in Generative AI and Virtual Worlds', *Competition Policy Brief*, 3, p. 7.

sections 2.3.1 and 2.3.2).[191] Downstream, they all have ready access to deployment routes for FMs in consumer and business products and services. These include business software offerings, search engines, operating systems and smart phones, where generative AI tools can be used or integrated.[192]

For example, some large digital firms already pre-install their models (e.g. search) or sell them as a bundle with other products (e.g. smart phones). Specifically:

- Microsoft has integrated its AI Companion 'Copilot' into its Microsoft 365 Apps.[193]
- Alphabet has started to pre-install its AI assistant 'Gemini' in its smartphones and provides AI generated overviews in Google Search.[194]
- Meta recently integrated its Meta AI assistant in WhatsApp.[195]

There are also some large digital companies which own platforms or marketplaces that provide access to AI models of competitors (e.g. Amazon Bedrock or Alphabet's Vertex AI).

This integration across the AI value chain gives rise to firm specific ecosystems of complementary products and services and may provide the respective firms with a significant competitive advantage.[196] It can lead to cost savings and improved efficiency due to the use of common components on the infrastructure level of the AI supply chain or improvements in the quality of the services provided due to data enabled learning. The users of these ecosystems benefit from increased convenience as they have access to 'one-stop shops',[197] exposure to new products and services, and a more seamless user experience as less interoperability issues are likely to appear when the integration of user data across services is centralised in the ecosystem.[198]

At the same time, vertical integration could contribute to higher barriers to entry for competitors if consumers (i) tend to stick to a default setting due to search and switching costs or a status-quo bias (**'consumer inertia'**),[199] or the **choice architecture** in the ecosystems pressures users to choose a specific AI model.[200] While integration of products and services is often a pro-competitive business practice that characterises healthy competition and drives

[191] ADLC (2024), 'opinion 24-A-05 on the competitive functioning of the generative artificial intelligence sector', 28 June, para. 284.
[192] ADLC (2024), 'opinion 24-A-05 on the competitive functioning of the generative artificial intelligence sector', 28 June, para. 287.
[193] Microsoft (2025), 'Set up Microsoft 365 Copilot – admin guide', accessed 2 May 2025.
[194] ADLC (2024), 'opinion 24-A-05 on the competitive functioning of the generative artificial intelligence sector', 28 June, para. 287; Google Search Central (2025), 'AI Overviews and your website', accessed 2 April 2025.
[195] WhatsApp (2025), 'About Meta AI', accessed 2 April 2025.
[196] Langus, G., Maier, N. and Muhamedrahimov, R. (2024), 'Balanced and Transparent Antitrust in the AI space' *Concurrences no.2-2024*, p. 13; Kowalski, K., Volpin, C. and Zombori, Z. (2024), 'Competition in Generative AI and Virtual Worlds', *Competition Policy Brief*, 3, p. 8.
[197] Langus, G., Maier, N. and Muhamedrahimov, R. (2024), 'Balanced and Transparent Antitrust in the AI space' *Concurrences no.2-2024*, pp. 13–14.
[198] Kowalski, K., Volpin, C. and Zombori, Z. (2024), 'Competition in Generative AI and Virtual Worlds', *Competition Policy Brief*, 3, p. 5.
[199] Wisnicky, B. (2022), 'Customer inertia fosters producto quality', *Journal of Behavioral and Experimental Economics*, **96**:101817, p. 1.
[200] CMA (2022), 'Online Choice Architecture: How digital design can harm competition and consumers´, discussion paper, April, pp. 16–23.

efficiency in the market,[201] and cannot be assumed to be anti-competitive a priori even when involving firms in a dominant position,[202] there can be specific instances when a dominant firm might use this as a strategy to foreclose competitors in a manner that harms consumers.

Indeed, the ADLC highlights the risk that large digital companies may have the ability to harm competition between AI developing firms by tying the sale of their product or service to the use of their own AI models (‘**tying and bundling practices’**) in the future.[203] Some large digital companies were previously fined for similar practices in the past. For instance, Microsoft was found to be foreclosing competitors by pre-installing its Internet Explorer on the user's Windows desktop and thereby using customer inertia to its own advantage.[204] Authorities are therefore likely to keep a close watch on this type of behaviour in AI markets.

Another related risk is that large digital firms may use their downstream platforms or marketplaces to promote the applications using their AI models (‘**self-preferencing**’).[205] This theory of harm has been a frequent concern in digital markets. For instance, in *Google and Alphabet v Commission* the European Commission found that ‘traffic from Google's general results pages was redirected […] in favour of Google's own comparison shopping service […]’.[206] In theory, Google could do the same for its AI-powered applications, promoting Gemini over HuggingChat or ChatGPT.[207] A similar concern could also arise for the platforms or marketplaces owned by large digital companies (e.g. Amazon Bedrock or Google Cloud), where certain AI models could be particularly promoted by large digital companies.

Additionally, large digital firms may also further increase the use of their own AI models by denying access to, or imposing discriminatory conditions for, accessing their products or services downstream (‘**refusal to grant access’** or ‘**discriminatory access**’). Specifically, they may refuse other AI application providers access to their products or services by not allowing them to be installed; or only allow access if the AI applications have been developed using the large digital companies' own closed-source AI models.[208] As a hypothetical example, Microsoft could require the use of OpenAI's GPT (Microsoft's partner) to develop AI-enabled applications for Microsoft Office. Given the benefits that would arise from the increased use of the incumbent's model downstream and the potential positive data feedback loops, this is another area of potential concern for competition authorities.[209]

[201] Langus, G., Maier, N. and Muhamedrahimov, R. (2024), ‘Balanced and Transparent Antitrust in the AI space’ *Concurrences no.2-2024*, p. 15.
[202] Bishop, S. and Walker, M. (2010), *The Competition of EC law: Concepts, Application and Measurement*, University edition, Sweet & Maxwell, pp. 6-064–6-084.
[203] ADLC (2024), ‘opinion 24-A-05 on the competitive functioning of the generative artificial intelligence sector’, 28 June, para. 287.
[204] Gore, D., van Rooijen, A. (2021), ‘Ex Post Assessment of European Competition Policy: The *Microsoft* cases’, *Draft report prepared for the 2021 Annual Conference of the GCLC*, p. 13.
[205] Carugati, C. (2023), ‘Competition in Generative Artificial Intelligence Foundation Models’, Working Paper, p. 11.
[206] *Google and Alphabet v Commission*, Case T-612/17, para. 244.
[207] Carugati, C. (2023), ‘Competition in Generative Artificial Intelligence Foundation Models’, Working Paper, p. 13.
[208] Carugati, C. (2023), ‘Competition in Generative Artificial Intelligence Foundation Models’, Working Paper, p. 14.
[209] Carugati, C. (2023), ‘Competition in Generative Artificial Intelligence Foundation Models’, Working Paper, p .14.

Potential mitigating factors

Large digital companies may be prevented from employing leveraging practices by the potentially disruptive impact of AI models downstream. Specifically, large digital companies will only be able to successfully engage in leveraging practices if they retain their position as dominant firms in downstream markets. In the past, the introduction of new technologies has led to companies with a previously very high market share quickly declining in importance. A prominent example of this is Nokia, which lost almost its entire user base within only a few years of the launch of smartphones.[210]

AI may similarly disrupt some of the markets where large digital companies currently have entrenched dominant positions.[211] For instance, Alphabet is vulnerable to consumers using LLMs (e.g. ChatGPT) or AI-powered voice assistants instead of Google Search to answer queries. Amazon, as an online marketplace, could also be replaced by effective AI agents recommending products in the future.[212] These companies may lose their dominant positions in these markets if other AI developing companies effectively create deployment routes downstream (such as OpenAI with ChatGPT).[213] Large digital companies may then need to refer to third-party AI models to remain competitive. For instance, in the future, Alphabet may be obliged to pre-install another AI assistant other than 'Gemini' in Google Search so that people keep using its search services.[214]

More generally, leveraging by a vertically integrated firm only makes economic sense if upstream benefits from foreclosing and self-preferencing are higher than the downstream benefits received from offering a diverse range of services that meets (growing) consumers' needs.[215] It may thus be that vertically integrated incumbents have incentives to grant access to as many AI models as possible in the downstream layer rather than self-preferencing. For example, Amazon is constantly expanding its Bedrock platform with new generative AI models as opposed to self-preferencing its own models.[216]

In any case, large digital companies' leveraging practices may also be regulated under sectoral provisions in certain jurisdictions.[217] As mentioned in section 3.2.2, large digital companies have been designated as gatekeepers in relation to their core platform services (CPS) under the EU's DMA. For these CPSs in particular, the EU can determine that applications using third-party AI models can be installed, and that no preferential treatment of

210 Zaman, R. (2020), 'Evolution of Smartphone as Disruptive Innovation', The Waves, 1 December, accessed 4 June 2025.
211 Hagiu, A. and Wright, J. (2025), 'Artificial Intelligence and competition policy', *International Journal of Industrial Organization*, 103134, p. 11.
212 Hagiu, A. and Wright, J. (2025), 'Artificial Intelligence and competition policy', *International Journal of Industrial Organization*, 103134, p. 12.
213 Hagiu, A. and Wright, J. (2025), 'Artificial Intelligence and competition policy', *International Journal of Industrial Organization*, 103134, p. 12.
214 ADLC (2024), 'opinion 24-A-05 on the competitive functioning of the generative artificial intelligence sector', 28 June, para. 287; Google Search Central (2025), 'AI Overviews and your website', accessed 2 April 2025.
215 See, for example, Dino, A. (2025), 'Survey finds consumers increasingly expect generative AI assistance in shopping', Lexpert, 19 March, accessed 28 May 2025. In businesses, nearly all organisations have multiple GenAI applications in production. The top five GenAI applications are being used by more than two-thirds of organisations. Techanalysis Survey Report (2024), 'The Intelligent Path Forward: GenAI in the Enterprise'.
216 AWS (2024), 'Supported foundation models in Amazon Bedrock', accessed 28 May 2025.
217 Ibid., p. 9.

gatekeepers' own or third-party services should be given. For instance, Google Android and Microsoft Windows already include services using third-party AI models.[218] In the UK, similar provisions apply under the DMCCA.[219] However, there remain outstanding questions in how far any regulatory provisions would need to be adapted to account for AI-specific features of a market. For example, it is unclear if the EU DMA app store obligations would apply in the same way to AI model app stores.[220] Gatekeeper services that are not classified as CPSs are also not covered by the EU DMA. [221]

The next section will discuss what powers of intervention are available to authorities to mitigate the potential risks identified in this section.

3.3 Questions for discussion

Box 3.1 Questions for discussion on tipping

- Are the current dynamics observed in the AI value chain signs of healthy competition or potential indicators of future competition concerns?
- Can competition be sustained at each layer of the AI value chain or is tipping likely to emerge in some of them?
- Which layers of the value chain are most contestable?

Box 3.2 Questions for discussion related to potential competition concerns

- What are the most significant barriers to entry, if any, in AI-related markets?
- Are any barriers in AI markets similar to those in other digital or non-digital markets? What is different?
- Is there a need to intervene in AI markets to protect competition?
- Is AI likely to disrupt current gatekeepers or more likely to entrench their position?

[218] Ibid., p. 9.
[219] CMA (2024), 'Digital Markets Guidance', guidance, 19 December, para. 4.23.
[220] Martens, B. (2024), 'Why artificial intelligence is creating fundamental challenges for competition policy?', *Policy Brief*, No. 16/2024, Bruegel, p. 10.
[221] Martens, B. (2024), 'Why artificial intelligence is creating fundamental challenges for competition policy?', *Policy Brief*, No. 16/2024, Bruegel, p. 9.

4 Do authorities have the power to regulate competition in AI?

The previous section identified several sources of competitive concern in AI markets—potential for market-tipping, foreclosure, discriminatory access, self-preferencing, tying and bundling—that may in some cases justify regulatory intervention. If such risks materialise, or are present already, and the authorities are willing to intervene, the next question is whether they have adequate powers to address these risks.

This section will review the extent to which any identified market failure could be addressed by traditional competition law (i.e. market inquiries, merger control and antitrust) or pre-existing regulation.

4.1 Are the traditional competition tools sufficient for AI markets?

4.1.1 Market inquiries

One tool in the standard toolkit of several competition authorities is the market inquiry. This is a specific competition tool that allows authorities to scrutinise whole markets or industries when competition seems weak or structural issues appear. Unlike normal antitrust cases targeting a specific cartel or abuse by one firm, market inquiries probe industry-wide features (concentration, barriers, pricing patterns, regulation, etc.) that may impede effective competition.

There are two main characteristics that make market inquiries powerful tools.

- First, they serve as a **preventive and diagnostic** competition tool. They identify issues in the market so that remedies or policy changes can be designed, while also acting as a silent deterrent to anti-competitive behaviour.
- Second, market inquiries also serve as a **signalling and fact-gathering exercise**. They can mobilise public debate, draw political attention to market failures, and build evidence for longer-term action.

Former EU Commissioner Margrethe Vestager noted that the Commission's sector inquiries (e.g. into energy) created 'political momentum' to strengthen enforcement and reform regulation.[222] Likewise, FTC Section 6(b) studies in the US (an analogous power) are used to obtain otherwise hard-to-get data on important sectors and can illuminate where policy action may be needed.[223]

To date, as discussed throughout this briefing paper, a number of authorities have deployed market inquiries to scrutinise AI-related markets. These include: the UK CMA, France's ADLC,

[222] Ashurst (2019), 'The use of sector inquiries in EU competition law', 24 October, accessed 28 May 2025.
[223] Lane, M. (2019), 'The FTC's 6(b) Study Authority: An Important Tool for Policymakers', DisCo, 9 April, accessed 28 May 2025.

Hungary's Gazdasági Versenyhivatal (GVH), the US FTC under Section 6(b), and the Competition Bureau of Canada.

Authorities can use their findings in various ways. In the EU, the Commission can launch formal antitrust investigations under Articles 101/102 TFEU if the inquiry uncovers suspected cartels or abuses. It may also propose regulatory or legislative changes to address cross-border issues not covered by competition law.[224] The CMA, in the UK, can impose direct remedies if it finds an 'Adverse Effect on Competition' (AEC) after a market investigation. Even without enforcement powers, authorities often make recommendations to regulators or the government—for example, by suggesting new rules or codes of conduct. Some of these actions have already been implemented with regards to AI markets, while in other cases, the competition authorities indicated that they will continue to observe the development of AI markets and intervene in a more informed manner due to the information gathered when needed (see section 4.2 for a further discussion).

4.1.2 Merger control

Another tool available to competition authorities is merger control.

Usually, one of the main points of focus for the authorities in merger reviews is to investigate whether the combined operation will increase prices. However, especially in those markets where price is not the main or only factor, authorities place considerable emphasis on non-price elements by investigating effects such as reduced innovation, diminished choice for consumers, or effects impacting privacy, data security, or the ease of consumers switching between platforms.[225] In the context of the AI sector, this means that authorities are often interested in the potential innovation that AI developer firms can bring to the market, their ability to choose between different compute power suppliers or between different distribution services, and, at the most downstream market, the choice that final consumers have in services where AI is implemented.

To date, merger control has been used by different authorities around the world to investigate the AI partnerships described in section 2.4.3, albeit with differing degrees of intervention.

The European Commission was involved in two different cases, Microsoft/Inflection and Nvidia/Run:ai, that showed the difficulties in capturing these kinds of agreements under the jurisdiction of EU Merger Regulation (EUMR).

- While seven different authorities referred the Microsoft/Inflection case to the Commission for a review, they all withdrew the referral after the Illumina/Grail judgment considerably limited the power related to Article 22 of the EUMR. The case was then investigated by the German competition authority which, however, found that it did not meet its jurisdictional threshold, i.e. Inflection did not have substantial operations in Germany.[226]

[224] Ashurst (2019), 'The use of sector inquiries in EU competition law', 24 October, accessed 28 May 2025.
[225] OECD (2025), 'Competition enforcement', accessed 30 May 2025.
[226] Peristerakis, N., Todisco, V., Keating, K. (2025), 'European Union: the Commission's evolving approach to digital mergers', in *Digital Markets Guide - Fourth Edition*, Global Competition Review.

- For the second case, Nvidia/Run:ai, the Commission accepted a referral from the Italian competition authority of a below-threshold transaction even for Italy and therefore raising the same concerns as those present in the Illumina/Grail judgment.[227] Following this investigation by the Commission—which concluded with an unconditional approval—Nvidia brought a lawsuit against the Commission for the investigation, criticising its use of Article 22 for being allegedly in contrast with the Illumina/Grail judgment.[228]

The CMA has been the most active authority in investigating AI partnerships through merger control tools. It opened inquiries into the five cases summarised in Table 4.1 below, selected from more than 90 partnerships that it identified (see section 2.4.3). This level of activity allows us to draw on a valuable basis for assessing the applicability of the merger control tool in AI markets—specifically, the limitations of the tool, as well as the extent to which it can address concerns raised regarding the risks of AI partnerships.

[227] Foo, Y.C. (2025), 'Nvidia takes EU antitrust regulators to court for probing AI startup Run:ai bid', Reuters, accessed 6 May 2025.
[228] Ibid.

Table 4.1 CMA merger inquiries in the AI sector

Case	Reason for clearance	Sources of influence/control	Turnover test or share test	Other insights from the CMA
Microsoft/OpenAI (partnership)	Cleared since there was no relevant merger situation due to no change of control. The CMA found that Microsoft already had a material influence on OpenAI, but it did not find a change in this control.	Assessed	**No conclusion**	The CMA identified the areas of potential competition concerns as the evolving market for the supply of accelerated compute, and the markets for development of FMs, distribution of FMs, and supply of FM-based services; however, no action was taken as a consequence of this investigation.
Microsoft/Mistral (partnership)	Cleared on the basis that there was no relevant merger situation since Microsoft did not acquire material influence over Mistral.	Assessed	**No conclusion**	No further insights provided by the CMA.
Amazon/Anthropic (partnership)	Cleared on the basis that there was no relevant merger situation as Anthropic did not have a UK turnover above £70m, nor did the parties together account for at least 25% of the supply of any description of goods or services in the UK.	Assessed	The tests were not met. The CMA said that it considered various descriptions of goods and services and various measures of supply, with no further details.	The CMA did not reach a conclusion on Amazon's influence on Anthropic, however it stated that agreements between an FM developer and a CSP for compute infrastructure and distribution may result in an acquisition of material influence.
Alphabet/Anthropic (partnership)	Cleared on the basis that there was no relevant merger situation since Alphabet did not acquire material influence over Anthropic.	Assessed	The CMA found that the turnover test was not met and **did not reach a conclusion on the share of supply**.	The CMA highlighted the overlap between the two parties in the supply of FMs and operations in the downstream AI application market. It also found a vertical relationship as Alphabet provides compute and distribution services to FM developers, including Anthropic. Finally, the CMA highlighted that Alphabet offers products and services in which FMs are or will be integrated.
Microsoft/Inflection (acqui-hire)	Cleared on the basis that it does not give rise to a realistic prospect of a substantial lessening of competition as a result of horizontal unilateral effects.	Assessed	The CMA found that the turnover test was not met, while the share of supply test was met. The CMA highlighted how it has a broad discretion in applying the latter test.	The CMA found overlap in the markets for: - development and supply of FMs globally; - development and supply of consumer chatbots globally. The CMA then presented two theories of harm on horizontal unilateral effects in these two markets.

Source: CMA (2025), 'Microsoft Corporation's partnership with OpenAI, Inc. Decision on relevant merger situation', full text decision, 15 April. CMA (2025), 'Microsoft Corporation's partnership with Mistral AI. Decision on relevant merger situation', full text decision, 21 May. CMA (2025), 'Amazon.com Inc.'s partnership with Anthropic PBC. Decision on relevant merger situation', full text decision, 17 October. CMA (2024), 'Alphabet Inc.'s partnership with Anthropic PBC. Decision on relevant merger situation', full text decision, 24 December. CMA (2024), 'Microsoft corporation's hiring of certain former employees of Inflection and its entry into associated arrangements with Inflection. Decision on relevant merger situation and substantial lessening of competition', full text decision, 24 October.

First, uncertainty remains on whether the CMA's use of merger control has highlighted that the cases assessed did not present competition concerns or whether it exposed this tool's limits in being effective in the context of AI partnerships. While able to start merger inquiries, the CMA found that four out of the five cases did not qualify as relevant merger situations. This was due either to (i) the absence of a change in control, or (ii) the failure of the share of supply and turnover tests. As part of these cases, the CMA provided clarity to businesses regarding its jurisdictional powers in the AI sector. [229] For instance, in the Microsoft/Inflection case, the CMA showed the ability to adopt a broad definition for a merger capturing within it the 'acqui-hire', i.e. where there is a transfer of employees from one firm to another that results in the creation of a relevant merger situation. In other cases, it gave insights both on the potential competition concerns that it identified and on the flexibility that it has in conducting the share of supply test. On one hand, when the CMA avoided reaching a conclusion on the competitive impact of the AI partnership in light of the transaction's failure to qualify under the merger regime, raises the question of whether the CMA was signalling a need for broader intervention powers since it found that the majority of the transactions fall outside its scope of intervention. [230]

Second, there is another potential limitation of merger control that was highlighted by the CMA's activity. Since the partnerships investigated have implications that go beyond the two involved parties and have an effect on multiple AI markets and their evolution, they should not be seen as 'an isolated series of bilateral deals, but [they represent] a new market structure where collaboration, rather than competition, is becoming the norm'.[231] Therefore, while the size of the single firm may be negligible according to the standard merger control thresholds, the size of the phenomenon is considerably larger. What is being observed is the creation of a market where a collaborative structure is taking the place of competition-based interaction. This interests the whole market by affecting the incentives that all the firms have. In the past, similar collaborative structure characterised highly innovative markets, such as the biotech industry in the 1990s. The matter therefore is not whether a degree of collaboration should be allowed, but what type of collaboration and whether the design of the agreements will foster innovation or hinder it.[232]

The European Commission and CMA's experiences with merger control in AI markets leave open questions regarding the effectiveness of this tool in the context of AI partnerships. Will the current application of merger control provide sufficient clarity until a conventional merger situation arises? Is there a need for a more in-depth investigation of the AI partnerships? Or, considering the potential systematic implications, is the use of an alternative regulatory instrument necessary to effectively analyse these collaborations?

[229] Herbert Smith Freehills (2024), 'Key themes emerging from the UK CMA's review of AI partnerships under its merger control rules', *Competition, Regulation and Trade eBulletin*, 21 November.
[230] Linklaters (2024), 'The rise of 'intelligent' partnerships', 6 September, accessed 28 May 2025.
[231] Groza, T. (2025), 'AI Partnerships Beyond Control Lessons from the OpenAI-Microsoft Saga', *Stanford Law School Blogs/CodeX*, 21 March.
[232] Ibid.

4.1.3 Antitrust

Competition authorities also have tools to address competition issues *ex post*, via antitrust rules that forbid abuses of dominant position (e.g. Article 102 TFEU in the EU) and anti-competitive agreements (Article 101 TFEU in the EU).

Antitrust enforcement in AI markets remains in its early stages, with only few precedents and investigations, which are either (i) a sign that AI markets do not yet raise competition concerns as there are no abuses of dominance, or (ii) a sign of authorities being slow to identify the concerns. In light of the scrutiny of the sector, discussed in sections 4.2 and 4.3, the latter seems unlikely at this stage. Ongoing investigations and antitrust precedents are cited below.

Nvidia

The ADLC has expressed concerns regarding the industry's reliance on Nvidia for AI chips, emphasising the potential for leveraging Nvidia's CUDA computing platform as the only environment fully compatible with common AI frameworks, or Nvidia's investments in AI-focused CSPs.[233] These concerns are part of a broader investigation into Nvidia's (potentially abusive) business practices, that was opened in July 2024 after dawn raids in July 2023 at their offices. The US DOJ[234] and China[235] have also launched antitrust investigations against Nvidia, suspecting behaviour aimed at entrenching its dominance in the AI chip market.

Alphabet (Google)

At the frontier between antitrust and copyright regulation, in March 2024, the ADLC has fined Alphabet (Google) for unauthorised content scraping from online news websites to train its Gemini AI chatbot.[236] The ADLC had investigated Google's behaviour after a complaint from an association of French publishers. The ADLC had pointed to multiple competition concerns in its preliminary assessment, including potential unfair trading conditions (imposed by Google on press agencies by refusing to negotiate and pay for the display of protected content) that could constitute an abuse of Google's dominant position in the French market for generalist search services. Google proposed commitments that were accepted and made binding in a decision of 2022, and the 2024 fine stemmed from Google's failed commitment that required to negotiate in good faith

[233] ADLC (2024), 'opinion 24-A-05 on the competitive functioning of the generative artificial intelligence sector', 28 June.
[234] Bloomberg (2024), 'Nvidia gets DOJ subpoena in escalating antitrust probe', September, accessed 28 May 2025.
[235] The New York Times (2024), 'China opens investigation into Nvidia over potential antitrust violations', December, accessed 28 May 2025.
[236] Decision 24-D-03 of 15 March 2024 regarding compliance with the commitments in Decision 22-D-13 of 21 June 2022 of the ADLC regarding practices implemented by Google in the press sector.

and with transparency with French news publishers over the use of their content and remuneration.[237]

While this briefing paper does not explore the topic of copyright in depth, this precedent demonstrates the existing (and growing) links between competition law and copyrights regulation. Copyright infringement litigation against AI developing firms is currently underway in several countries, reflecting global concerns over the unauthorised use of copyrighted materials in training AI models. However, these are lengthy processes with unclear outcomes that might also be insufficient to address some of the other concerns associated with delayed intervention (e.g. markets tipping in the meantime). ADLC's fine against Google shows that there is scope that copyright cases may be pursued under competition law in the future, although there is uncertainty around how this will eventually unfold.

The effectiveness of antitrust enforcement in the AI sector is debated. In particular, the concern is that in fast-moving digital sectors, *ex post* antitrust action is often too slow or too weak to restore competition once dominant positions are entrenched (e.g. due to market tipping) and may also struggle to design appropriate remedies. This underpins the ongoing debate on the appropriate timing of intervention, explored further in Section 5.

4.2 To what extent do current regulations address AI competition risks?

At the European level, there is no competition law tailored specifically to intervention in AI markets. There has been a range of regulations enacted, however, that may mitigate AI competition risk either directly through contractual obligations or structural remedies, or indirectly through data regulations and wider investigative powers. These include:

- The **DMA**, which establishes obligations for companies that may abuse their market power in digital sectors.[238]
- The **AI Act**, which sets out a risk-based framework to regulate AI in the EU.[239]
- The **General Data Protection Regulation (GDPR)**, which provides comprehensive rules on the collection, use, and protection of personal data within the EU.[240]
- The **Data Act**, which aims to promote fair access to and use of data across sectors in the EU.[241]

[237] Decision 22-D-13 of 21 June 2022 regarding practices implemented in the press sector.
[238] EU (2025), 'The Digital Markets Act', accessed 28 May 2025.
[239] European Council (2024), 'Artificial Intelligence Act', accessed 20 May 2025.
[240] European Parliament and Council (2016), 'General Data Protection Regulation', *Official Journal of the European Union*, legislative act, 4 May.
[241] European Commission (2024), 'Data Act', accessed 20 May 2025; European Commission (2024), 'Data Act explained', accessed 28 May 2025.

- The **Directive on copyright in the Digital Single Market (DSM)**, which seeks to harmonise and update copyright regulations across the EU.[242]

Of these, the GDPR, Data Act and the DSM might indirectly mitigate competition risks in AI markets by regulating companies' access to data. Specifically, the regulations may influence how AI systems handle personal data, what data AI system have access to, and the conditions under which data can be shared or used. As such, they may mitigate large digital firms gaining a significant competitive advantage from their preferential access to proprietary datasets (see section 3.2.2).[243]

Additionally, the EU AI Act is expected to prevent competition concerns more directly by significantly extending competition agencies' investigative powers.[244] Usually, competition authorities can request information from companies only where they have evidence of a competition law breach. Under the EU AI Act, market surveillance agencies have the power to conduct compliance checks independently, and are expected to provide yearly reports to competition authorities. They also have significantly more access to sensitive information than competition authorities generally do.[245] This means that any competition concern in relation to AI developing firms will likely be noticed early.

The main active regulations that have a potential remit for intervention in AI markets are the DMA in the EU and the DMCCA in the UK. As explained in section 3.2, both of these regulations impose restrictions on gatekeepers' activities in relation to CPSs, e.g. Alphabet's activities in relation to Google Search.[246] The EU High-Level Group has explicitly clarified that the provisions under the DMA can apply to AI systems 'to the extent that [they] are embedded into designated CPSs'.[247] The CMA has similarly indicated in the Updated Foundation Models paper that they may consider investigating 'digital activities that are critical access points or routes to market for FM development'.[248]

The gatekeeper obligations in the DMA and DMCCA might prove to be sufficient to regulate AI developing firms as of now. This is because the major concerns at this stage are related to firms that meet the designation criteria, i.e. Alphabet, Amazon, Apple, Meta and Microsoft.[249] These firms have already been designated as gatekeepers for some platforms

[242] European Parliament and Council (2019), 'Directive on copyright and related rights in the Digital Single Market', *Official Journal of the European Union*, legislative act, 17 April; International Federation of Library Association and Institutions (2019), 'European Directive on Copyright in the Digital Single Market: what it is about and why libraries should care', 18 June.
[243] ADLC (2024), 'opinion 24-A-05 on the competitive functioning of the generative artificial intelligence sector', 28 June, paras 197–199.
[244] Schrepel, T. (2025), 'Decoding the AI Act: Implications for Competition Law and Market Developments', *Journal of Competition Law &* Antitrust, **00**, p. 3.
[245] Schrepel, T. (2025), 'Decoding the AI Act: Implications for Competition Law and Market Developments', *Journal of Competition Law &* Antitrust, **00**, pp. 3–4.
[246] EU (2025), 'Gatekeepers', accessed 28 May 2025.
[247] European Commission (2024), 'High-Level Group for the Digital Markets Act Public Statement on Artificial Intelligence', 22 May, accessed 28 May 2025.
[248] CMA (2024), 'AI Foundation Models - Update Paper', 11 April, para. 41.
[249] CMA (2024), 'AI Foundation Models', technical update report, 16 April, para. 5.3; Competition Bureau Canada (2024), 'Artificial intelligence and competition – Discussion Paper', 20 March, pp. 17–18; AdC (2023), 'Competition and Generative Artificial Intelligence', Issues Paper, November, p. 27.

in which AI systems may be embedded by the EU.[250] The UK has similarly started its first investigation to potentially assign 'Strategic Market Status' (i.e. the UK's 'gatekeeper' equivalent) to Alphabet in relation to Google Search.[251]

As explained in section 3.2, conduct-based rules in the DMA therefore already allow the EU to ensure access to key inputs such as cloud computing and data, and mitigate some of the concerns that arise from the vertical integration of large digital companies. For instance, the following obligations apply to gatekeepers under this regulation:

- they cannot use technical or other means to restrict switching;[252]
- they need to provide competitors with access to certain query data, which could be extended to generative AI;[253]
- they are barred from ranking or integrating their in-house AI services ahead of rivals.[254]

On the more extreme end of intervention, there also exist options of structural remedies that limit the lines of business in which a single firm can operate. The objective of this approach would be to prevent firms with power in upstream and downstream markets from leveraging that market power to exercise self-preferencing or foreclosure.[255] For instance, this could mean mandating firms that provide both cloud computing and AI FMs to separate those divisions, thereby removing the ability and incentive to favour its own AI services over competitors using its cloud platform. Such structural separation is a more extreme form of intervention than conduct-based rules, as it reshapes the market structure itself to pre-empt risks in vertically integrated AI value chains. However, such a type of intervention would prevent the efficiency gains that are characteristic to vertically integrated firms, and on balance could lead to a less competitive outcome.

In this context, an issue may be that the restrictions of the DMA and DMCCA only cover activities of gatekeepers in CPSs.[256] These do not include all services in which AI models may be embedded. However, both regulations have the ability to expand the set of designated firms in case new players emerge that meet the criteria and there is potential, over time, to expand the focus of the services that are addressed by the regulations to specifically capture AI as a service. For instance, cloud services are included as a CPS, although no gatekeepers have been designated yet for them.[257]

[250] EU (2025), 'Gatekeepers', accessed 28 May 2025.
[251] CMA (2025), 'Notice under Section 11(1) of the Digital Markets, Competition and Consumers Act 2024 (The Act)', investigation notice.
[252] Autoriteit Consument & Market (2023), 'Market Study Cloud services', 5 September, p. 70.
[253] Martens, B. (2024), 'Why artificial intelligence is creating fundamental challenges for competition policy', Bruegel.
[254] European Commission (2022), 'Digital Services Package: Commission welcomes the adoption by the European Parliament of the EU's new rulebook for digital services', press release, 5 July.
[255] Cath, C. (2024), 'Is "More Clouds" the Future We Want? A Dispatch from the FTC AI Tech Summit', Tech Policy Press.
[256] Kowalski, K., Volpin, C. and Zombori, Z. (2024), 'Competition in Generative AI and Virtual Worlds', *Competition Policy Brief*, 3, p. 9.
[257] Autoriteit Consument & Market (2023), 'Market Study Cloud services', 5 September, p. 69.

While the debate on the need for specific AI competition regulation is in the incipient phase, this might indicate that future regulation will not need to be developed from scratch but instead can be added to existing ones.

4.3 Questions for discussion

Box 4.1 Questions for discussion on legal powers and sector regulation

- Are the recent partnerships analysed by the CMA truly unproblematic or did the CMA just not have the power to intervene? What does this mean for the merger regime when it comes to AI firms?
- Is merger control able to address the structural changes presented in the AI market and its systematic implications?
- Is there a case for developing separate AI competition regulation or are any market failures sufficiently addressed by existing competition law and regulation?
- Is there a need to extend current regulations to address AI specific competition concerns?

5 The right time to intervene

In this section we consider the crucial and often contentious dimension of timing: at what stage, and based on what evidence of likely tangible harm, should authorities consider intervening in fast-developing markets? This exploration considers the delicate balance between pre-empting anti-competitive outcomes and avoiding the premature stifling of innovation and draws lessons from intervention in two other markets (aviation and GDPR).

5.1 Intervening too early or too late?

While some argue that the experience with digital platforms suggests early intervention might prevent market concentration,[258] how should regulators determine the optimal moment and mode of intervention in AI markets? In light of the ongoing debate related to the existence—or absence—of meaningful barriers to entry and the dynamic developments in the market, there is a concern that intervening too early will be detrimental and lead to regulators picking winners or stymying innovation.[259] Should regulation await better-defined market structures, or would doing so risk repeating past mistakes made during the rise of digital markets?

The challenge lies in balancing the risks of delayed action against premature regulation that could distort the market evolving efficiently.[260] It is an open question whether AI-driven markets have already reached a stage where regulatory designation criteria can be meaningfully applied. The complexity of defining AI and its rapidly evolving capabilities further complicates this assessment.

There can be benefits associated with intervening early, provided the right conditions are met. For instance, it may be a relevant policy objective to prevent the expansion of excessive control over essential inputs, such as computing infrastructure, and it may be least costly to do so before it expands into the AI area. Although new entrants and smaller firms may benefit from the frameworks and tools developed by innovating incumbents, training a state-of-the-art AI model already requires months of GPU time and special access to advanced node fabrication, suggesting the potential for a significant first-mover advantage.[261] The concern is a potential future scenario in which a few tech companies dominate AI markets and shape regulation around them.

[258] Cabral, L. et al. (2021), 'The EU digital markets act', *Publications Office of the European Union*. Hua, S.S. and Belfield, H. (2023), 'Effective Enforceability of EU Competition Law Under AI Development Scenarios: a Framework for Anticipatory Governance', *Proceedings of the 2023 AAAI/ACM Conference on AI, Ethics, and Society*, pp. 596–605.
[259] Bostoen, F. and van der Veer, A. (2024), 'Regulating Competition in Generative AI: A Matter of Trajectory, Timing and Tools', working paper.
[260] Lancieri, F. and Pereira Neto, C.M.S. (2020), 'Designing remedies for digital markets: The interplay between antitrust and regulation', *Journal of Competition Law & Economics*, **18**:3, pp. 613–669.
[261] Krishnakumar, A. et al. (2023). 'Domain-specific architectures: Research problems and promising approaches', *ACM Transactions on Embedded Computing Systems*, **22**:2, pp. 1–26.

Although some commentators have argued in favour of regulation, others have highlighted the risks of intervening too early. Bostoen and van der Veer (2024) argue that generative AI markets are still in a nascent stage, lacking clear market failures that justify immediate regulatory intervention.[262] Moreover, it is not always necessarily clear how developments in AI would facilitate anticompetitive conduct. For example, Ittoo and Petit (2017) find that current AI technologies may have inherent limitations that naturally mitigate competitive risks, such as algorithmic collusion.[263] Historical evidence from previous technological waves suggests that adaptive regulation may help AI innovators thrive, with more targeted intervention to follow only if specific market failures arise.[264]

To consider where the right regulatory balance lies, it is relevant to consider the impact that similar regulations have had on the functioning of markets. Since the DMA was intended to address competition issues in digital markets, it is a prime candidate to draw learnings from. Again, opinions are divided. Some argue that the DMA may effectively regulate data-driven, multi-sided platforms due to their ex-ante approach to addressing gatekeeper power and their aim to balance efficiencies with competitive fairness.[265] However, others strongly caution against such early intervention. Notably, Davies et al. (2022) criticise the DMA's approach as contradicting evidence-based policy principles, arguing that excessive regulation risks stifling innovation especially in nascent markets.[266] This is because the underlying economics of different platform services may be distinct, making uniform regulatory obligations potentially harmful, or less effective. This criticism reflects broader concerns that premature intervention—without appropriate tailoring—could potentially disrupt natural market evolution and technological innovation.

5.2 Lessons from interventions in other sectors

The timing and scope of regulation are essential. Historical case studies suggest that, even if some market conditions may justify intervention, a trade-off must be made between acting too late—letting dominant firms entrench power—and acting too early—imposing burdens that could stifle innovation in a nascent industry.

For example, academic research on early regulatory attempts in commercial aviation shows the potential for harmful effects. Early regulatory attempts in commercial aviation had specific objectives related to fares and route allocation, in addition to safety.[267] A few decades later, deregulation removed control over fares and market entry, while

[262] Bostoen, F. and van der Veer, A. (2024), 'Regulating Competition in Generative AI: A Matter of Trajectory, Timing and Tools', working paper.
[263] Ittoo, A. and Petit, N. (2017), 'Algorithmic pricing agents and tacit collusion: a technological perspective', in H. Jacquemin and A. de Streel (eds), *L'intelligence artificielle et le droit*, Larcier, pp. 241–256.
[264] Kulothungan, V., Mohan, P.R. and Gupta, D. (2025), 'AI Regulation and Capitalist Growth: Balancing Innovation, Ethics, and Global Governance', working paper.
[265] Cabral, L. et al. (2021), 'The EU digital markets act', *Publications Office of the European Union.*
[266] Davies, J. et al. (2022), 'A missed opportunity: The European union's new powers over digital platforms', *The Antitrust Bulletin*, **67**:4, pp. 504–521.
[267] Pitt, I.L. and Norsworthy, J.R. (1999), 'A Brief Review of Airline Regulation', *Economics of the US Commercial Airline Industry: Productivity, Technology and Deregulation*, pp. 67–96.

maintaining safety standards.[268] In the US, any domestically-owned airline deemed 'fit, willing, and able' was allowed to operate on any domestic route. Elsewhere, several state-owned assets were transferred to the private sector.[269] Borenstein and Rose (2007) find that lower fares and improved operational efficiency followed since deregulation. Airlines faced new competitive pressures that spurred experimentation in pricing, loyalty programmes, and flight scheduling. As such, the evidence on the impact of deregulation shows that initial regulation may have had harmful effects.[270]

Valuable lessons can also be learned from the implementation of GDPR. Despite its purpose to protect consumer privacy, it may have inadvertently led to higher market concentration. In particular, the need to receive user consent to use their data imposed additional costs for data collection and management.[271] Larger firms were better suited to absorb larger costs, even for non-compliance penalties.[272] Moreover, Gal and Aviv (2020) find that these fixed costs led to disproportionately higher per-user compliance costs for smaller firms because, unlike larger competitors, they could not match a unified consent framework across diverse products and millions of users. Some academic research associates GDPR's impact on data access with a 15% decrease in website utilisation of third-party web-service providers for EU users, and a 17% increase in concentration in that upstream market.[273] The experience with GDPR highlights how early intervention through strict regulation can inadvertently favour established, resource-rich firms while hampering competition and innovation.[274]

A similar risk could arise in the context of AI if there was premature intervention targeted at potential but unlikely competition risks. If regulation introduces compliance costs that are high and disincentivises investment and innovation, European firms—particularly newer or smaller players—could be disadvantaged relative to global incumbents better positioned to adapt. This would effectively tilt the playing field early on, entrenching those who already hold a foothold in the AI market and undermining Europe's competitiveness.

5.3 Geopolitics and the timing of intervention

As discussed in section 2.4.4, the geopolitical space within which AI is developing is fluid, strategic and increasingly contested. Interventions aiming for objectives other than the preservation of competition—most notably national security, technological sovereignty,

[268] Gowrisankaran, G. (2002), 'Competition and regulation in the airline industry', *Federal Reserve Bank of San Francisco*.
[269] Borenstein, S. and Rose, N.L. (2014), 'How airline markets work… or do they? Regulatory reform in the airline industry', in *Economic regulation and its reform: What have we learned?*, pp. 63–135.
[270] Button, K.J. (1996), 'Aviation deregulation in the European Union: Do actors learn in the regulation game?', *Contemporary Economic Policy*, **14**:1, pp. 70–80.
[271] Gal, M.S. and Aviv, O. (2020). 'The competitive effects of the GDPR', *Journal of Competition Law & Economics*, **16**:3, pp. 349–391.
[272] Frey, C.B. and Presidente, G. (2024), 'Privacy regulation and firm performance: Estimating the GDPR effect globally', *Economic Inquiry*, **62**:3, pp. 1074–1089.
[273] Johnson, G.A., Shriver, S.K. and Goldberg, S.G. (2023), 'Privacy and market concentration: intended and unintended consequences of the GDPR', *Management Science*, **69**:10, pp. 5695–5721.
[274] Bessen, J. et al. (2020), 'GDPR and the Importance of Data to AI Startups', *Boston University School of Law*, working paper.

and the resilience of critical supply chains—may exert a more immediate and powerful impact on AI markets than traditional competition policy and competition regulation. For example, a recalibration of political alliances between Europe and the US or a further erosion of trust between existing partners could prompt a variety of retaliation measures such as export controls on AI inputs or procurement bans on AI-enabled products from rival jurisdictions which would supersede the competition aims entirely.

In this context, there are growing calls for Europe to accelerate its efforts to adopt a more active industrial policy with regards to AI markets to develop and secure its own critical infrastructure while continuing the push for responsible governance frameworks to avoid being 'swayed by external influences'.[275] Initiatives such as the proposed 'EuroStack'—a federated cloud-to-edge infrastructure under European jurisdiction—illustrate the emerging policy mix: supply-side investment to be paired with stringent rules on trustworthy AI, with the dual aim of safeguarding strategic autonomy and stimulating market rivalry.[276] If implemented at scale, these efforts could intensify competition in upstream compute and cloud services, and strengthen Europe's bargaining position vis-à-vis large non-European providers. However, the overall effect on competition is uncertain. In a cooperative scenario, EuroStack could complement global value chains and enhance contestability. In a more adversarial scenario—marked by retaliatory restrictions and fragmented standards—it could entrench regional blocs and limit market entry.

The geopolitical uncertainties highlight that any adjustments to EU competition policy and regulation in AI must be calibrated to the evolving geopolitical context. Intervention that is premature or misaligned with broader industrial and security objectives risks either stifling nascent European players or proving ineffective if measures primarily driven by considerations on sovereignty are adopted at the same time.

5.4 Questions for discussion

Box 5.1 Questions for discussion on the timing of intervention

- How to identify the right time for intervention to protect competition?
- What are the most significant costs that regulation might impose on AI?
- Is more competition in AI markets going to help or hinder the solution to other non-competition market failures?
- What roles should competition regulation play given the AI geopolitics?

[275] Csernatoni, R. (2025), 'The EU's AI power play: between deregulation and innovation', *Carnegie Europe*, May, p. 2; Letta, E. (2004), 'Much More Than a Market – Speed, Security, Solidarity: Empowering the Single Market to Deliver a Sustainable Future and Prosperity for all EU Citizens', April; Draghi, M. (2024), 'The future of European competitiveness – In-depth analysis and recommendations', September.
[276] Bria, F., Timmers, P., and Gernone, F. (2025), 'EuroStack – A European Alternative for Digital Sovereignty', Bertelsmann Stiftung. Gütersloh, February.

6 Questions for discussion

Box 6.1 Questions for discussion on tipping

- Are the current dynamics observed in the AI value chain signs of healthy competition or potential indicators of future competition concerns?
- Can competition be sustained at each layer of the value chain or is tipping likely to emerge in some of them?
- Which layers of the value chain are most contestable?

Box 6.2 Questions for discussion related to potential competition concerns

- What are the most significant barriers to entry, if any, in AI-related markets?
- Are any barriers in AI markets similar to those in other digital or non-digital markets? What is different?
- Is there a need to intervene in AI markets to protect competition?
- Is AI likely to disrupt current gatekeepers or more likely to entrench their position?

Box 6.3 Questions for discussion on legal powers and sector regulation

- Are the recent partnerships analysed by the CMA truly unproblematic or did the CMA just not have the power to intervene? What does this mean for the merger regime when it comes to AI firms?
- Is merger control able to address the structural changes presented in the AI market and its systematic implications?
- Is there a case for developing separate AI competition regulation or are any market failures sufficiently addressed by existing competition law and regulation?
- Is there a need to extend current regulations to address AI specific competition concerns?

Box 6.4 Questions for discussion on the timing of intervention

- How to identify the right time for intervention to protect competition?
- What are the most significant costs that regulation might impose on AI?
- Is more competition in AI markets going to help or hinder the solution to other non-competition market failures?
- What roles should competition regulation play given the AI geopolitics?

Contact

Andreea Antuca
Senior Consultant
+44 (0) 1865 253205
andreea.antuca@oxera.com

oxera.com