Hard facts. Clear stories.

Copenhagen Economics

GENERATIVE ARTIFICIAL INTELLIGENCE: THE COMPETITIVE LANDSCAPE

White paper

PREPARED FOR: CCIA EUROPE
FEBRUARY 2024

AUTHORS
Paulo Abecasis
Dr Federico De Michiel
Dr Bruno Basalisco
Tuomas Haanperä
Julie Iskandar

This white paper is an independent report funded by the Computer & Communications Industry Association (CCIA Europe). The views and conclusions in this report are our own and do not represent the views of CCIA Europe.

ABOUT COPENHAGEN ECONOMICS
Copenhagen Economics is an expert-driven consulting company built on a deep knowledge of applied economics, and one of the leading economics firms in Europe.

We believe sound economic analysis can equip decision makers with hard facts and clear stories to make better choices for the benefit of society. We are committed to delivering compelling and pragmatic economics solutions with a creative and candid approach.

A brief note on consultancy research
As is standard in our field of professional services, research is designed so that:

- the client chooses the research question;
- we analyse and address the question to the best of our knowledge;
- findings and conclusions are our own.

The independence of our professional services is ensured via a diversified portfolio of business, spanning public sector and private clients across industries. For further information, see www.copenhageneconomics.com. We remain available for and appreciate any questions or comments.

TABLE OF CONTENTS

Copenhagen Economics

EXECUTIVE SUMMARY

The recent success of generative artificial intelligence (GenAI) tools capable of generating human-like content has shaken digital markets and started a new technological race. GenAI is powered by foundation models: large and complex AI models trained on vast amounts of data. Its impact is estimated to boost global GDP by 7% in the next decade according to a recent study.[1] Competition authorities across Europe have taken a forward-looking approach to understanding the market in its nascent stage and ensuring competition and innovation are preserved both presently and in the long run.

This white paper provides a preliminary review of current developments in the GenAI space and the emerging implications for competition. Further research is needed to shed additional light on the evolving competitive situation in GenAI, as well as any case-specific matters, particularly as this is a fast-evolving sector.

We have reviewed the research literature and identified three potential sources of competition concerns affecting GenAI markets:

1. limited access to inputs including data, computing power, hardware, and human talent may raise barriers to entry;
2. partnerships between large firms and new players may stifle potential competition;
3. leveraging behaviours by large firms may hamper competition in GenAI and other markets.

Our assessment of current market developments suggests that there are no evident signs of immediate competition concerns, with a number of new entrants present with diversified products and business models. It remains important to monitor and pre-empt future foreclosure concerns, which can be assessed under Article 102 or the Digital Markets Act in Europe.

At this stage, the market appears dynamic, with no or little sign of insurmountable barriers to entry. The number of foundation models launched is high and continues to increase with one third of more than 250 models available launched since August 2023 according to a database kept by Stanford University.[2] There are many open solutions available to non-integrated GenAI developers. New firms with different business models and degrees of specialisation compete head-to-head or even outperform established firms according to well-known rankings. More than 13 firms in GenAI have already achieved the coveted unicorn status.[3] Among those, OpenAI leads the GenAI race, while the recent European startup Mistral already established itself as a key competitor. A large influx of venture capital investments, with an almost fivefold annual increase to EUR 20 bn in 2023, allows startups to access costly inputs (thus reducing any barriers to entry) and signals that investors are confident in GenAI markets' competitiveness. Models are becoming increasingly smarter and less computationally demanding. However, there is uncertainty about whether inputs may become less accessible in the future. Some European GenAI startups face challenges in growing, and highlight regulatory costs (e.g., the AI Act).

Depending on their design, partnerships between large cloud providers and AI startups may give rise to competition concerns. Vertical collaborations between large digital

[1] Goldman Sachs (2023a).

[2] Stanford University (2024a).

[3] Unicorn status refers to a milestone in a startup's development when it reaches a market valuation of at least USD one billion.

players and GenAI startups are common and generate efficiencies, as they allow AI startups to access highly specialised hardware and computing power, additional support and investments.

However, if a larger partner uses its market power to exercise decisive control over a startup or gain privileged or exclusive access to its technology, this may harm competition across the value chain or may ultimately remove or dampen potential competition from the startup itself in the Gen AI or related markets. Any competitive implications will depend on the nature of the assets and capabilities brought together by the agreement, and its specific design. We find that partnerships are less likely to create competition concerns if there are a) no/limited exclusivity conditions, either in supply or distribution, and b) limited privileged access to the startup's valuable technological assets.

Integration of GenAI solutions into the existing services of large digital players can give rise to foreclosure. GenAI applications are most valuable when used together with existing services or products. This integration of GenAI into existing services can boost innovation in adjacent markets and potentially increase the competition faced by established players. However, this integration can also give rise to potential anti-competitive foreclosure via practices such as tying, bundling, or self-preferencing. Such practices should be monitored and assessed under Article 102 or the Digital Markets Act in Europe.

POTENTIAL COMPETITION CONCERNS

Based on a review of the existing literature, we identify three potential competition concerns that could apply to GenAI

Limited access to resources such as data, computing power, hardware, and talent

Contractual agreements between large firms (mostly cloud providers) and new players may stifle the emergence of new competitors in various markets

Adopting business practices that leverage market power into new markets or exploit existing power

CURRENT MARKET DYNAMICS

The market is dynamic, with new players emerging and large venture capital investments supporting open solutions and innovation in computational efficiency

Efficiency-enhancing partnerships are more likely to raise concerns when a large tech company has: control, benefits from strict exclusivity, or access to strategic technology of the start-up

The integration of GenAI into existing products is still nascent, and the risk of leveraging is captured by the DMA (if captured) or Article 102

1 GENERATIVE AI ATTRACTS ATTENTION FROM COMPETITION AUTHORITIES ACROSS EUROPE

1. The launch of the conversational chatbot ChatGPT by OpenAI in November 2022 attracted public attention to a new subfield of artificial intelligence, generative artificial intelligence (GenAI).[4] This began a technological race with a range of new models and applications announced in quick succession since then by established players and startups alike, such as Claude (Anthropic), LLaMA (Meta), various versions of Mistral, Gemini and Gemma (Google),[5] Mid-Journey, BLOOM (Hugging Face), Aleph Alpha, Titan (Amazon), Cohere, Inflection and Stability AI.

2. GenAI creates novel output seemingly indistinguishable from human-generated content, such as texts, images, videos, audio, and code based on users' inputs (e.g., in the form of text prompts).[6]

3. This transformative technology has the potential to revolutionise digital markets and change their competitive landscapes. More traditional industries are also looking at ways to use GenAI in their operations, from improving customer support to helping in drug discovery.[7] According to a study by Goldman Sachs, GenAI could boost global GDP by 7% within the next decade.[8]

4. User-facing GenAI applications such as OpenAI ChatGPT, Google Bard (now Gemini), Mistral Le Chat, or Microsoft Copilot are typically powered by underlying foundation models, such as OpenAI's GPT versions or Google's Gemini.

5. Foundation models are trained on large amounts of data and often can produce a variety of outputs across different domains and modalities; see Box 1Box 1 for a brief description of the training process. Large language models (LLMs) are the best-known foundation models and focus on natural language processing. They are generally used for text generation.[9]

---

[4] ChatGPT has set a record for having the fastest-growing user base in history for a consumer application, gaining 1 million users in just five days, and reaching 100 million monthly active users just two months after launch. See *Reuters* (2023a).

[5] Google (2024a).

[6] Multimodal foundation models can provide the backbone for various different user-facing applications producing text, images, codes etc.

[7] *Bloomberg* (2023).

[8] Goldman Sachs (2023a).

[9] Some examples of foundation models include GPT-4, DALL-E and Emu Video. GPT-4 is the multimodal model that powers ChatGPT and can generate coherent and fluent texts on various topics. It can also answer questions, perform calculations, and execute commands. GPT-4 is based on transformer architecture and has been pre-trained on a large corpus of web texts. DALL-E is a vision and language model that can generate realistic images from text descriptions, such as "a pentagon made of Cheese" or "an armchair in the shape of an avocado". It can also manipulate and combine images based on text instructions. DALL-E is a combination of GPT and a variational autoencoder and has been pre-trained on a large dataset of image–text pairs. Emu Video is a text-to-video model that can generate high-quality videos from texts or image prompts such as "a young couple walking in heavy rain" or "a teddy bear painting a portrait". It was developed based on previous text-to-images synthesis used with image generators (such as DALL-E).

Box 1 Steps to train and deploy GenAI models

Once a model is designed (e.g., choice of model architecture and number of parameters) it needs to be trained on existing data. Training is usually divided into two phases: pre-training and fine-tuning.

In the pre-training, the model is fed a vast amount of data, which it then uses to create its "knowledge" represented by billions of parameters and weights. Usually, pre-training is done on a large corpus of publicly available unstructured data (e.g., web crawling, books, news articles, or other openly available datasets). Proprietary data can also be used. This process is the most intensive in terms of computing power and can take several days or weeks.

Fine-tuning is the process of adapting the pre-trained model to a specific task or domain using supervised or semi-supervised methods, often on smaller and more specific sets of high-quality and domain-specific data. Fine-tuning may also include an alignment process to ensure that the model's behaviour aligns with the desired objectives and values of the users (e.g., to ensure non-harmful content in the answers). The fine-tuning phase also requires computing power, but arguably less than in the pre-training phase.

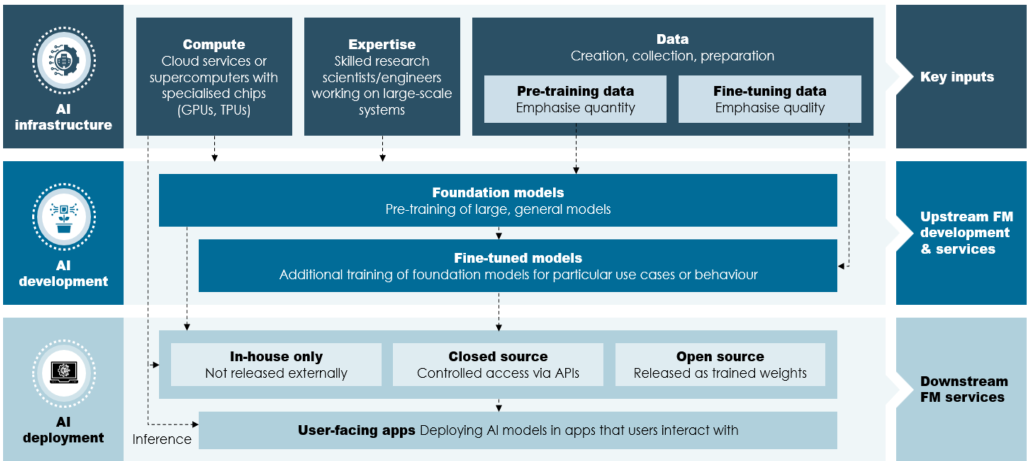
Finally, when the model is developed, it can make inferences, in other words, generate outputs or predictions based on new inputs or queries (i.e., prompts). While this phase requires the least computing power, it is executed every time the model is used and, therefore, it can be considered a (non-negligible) marginal cost.[10]

Source: Copenhagen Economics, based on CMA (2023a).

6. Many different firms are active in the GenAI value chain with different levels of vertical integration and openness. The value chain can be divided into the following three levels: AI infrastructure, AI development, and AI deployment; see Figure **1** based on the UK Competition and Markets Authority (the "CMA") AI Foundation Models Initial (CMA, 2023a).

[10] For example, when using the Mistral AI model based on its openly available API a single query, such as "Describe in detail the physics behind a hypothetical time machine. Please provide your sources", in its most advanced "mistral-medium" model required the usage of 898 tokens and cost €0.01.

Figure 1
GenAI value chain



Source: Copenhagen Economics, based on CMA (2023a).

7. The top level, AI infrastructure, includes the inputs necessary to develop and train foundation models: expertise from developers with specific Gen AI know-how (talent); large quantities of data that can be of different types; and computing power to process the data. Computers used for the development of foundational models are typically powered by Graphics Processing Units (GPUs), which are in short supply. Access to these inputs may require a significant amount of funds.[11]

8. The second level is AI development. This includes firms that develop and fine-tune[12] foundation models. Suppliers at this level of the chain are the core providers of GenAI models. A specific foundation model can be used to develop multiple fine-tuned models. Most foundation model developers produce fine-tuned versions of their own models, and some firms focus on developing fine-tuned models based on foundation models developed by others.[13]

9. Finally, at the bottom of the value chain is AI deployment. This is the stage where access to GenAI models is made available to end users. Distribution can occur via stand-alone applications or interfaces, such as the chatbots discussed above, or integrations into applications, such as Microsoft Copilot's integration on its 365 productivity suite.

10. AI developers can use various deployment and integration strategies: they have the option to offer the GenAI model as a stand-alone product or incorporate it as features to enhance their existing products (e.g., the Adobe Firefly model for generating and modifying images is offered in its Creative Cloud). Sometimes, features of GenAI models can be expanded through plug-ins developed by third parties, thus allowing for further enhanced task-specific customisation.

11. All foundational models have several components, such as model weights, training, and inference code,[14] training datasets, etc. Developers can choose to release publicly or keep private any combination of these components and often offer different forms of access (e.g., hosting access, API access, downloading) to competitors or customers.[15]

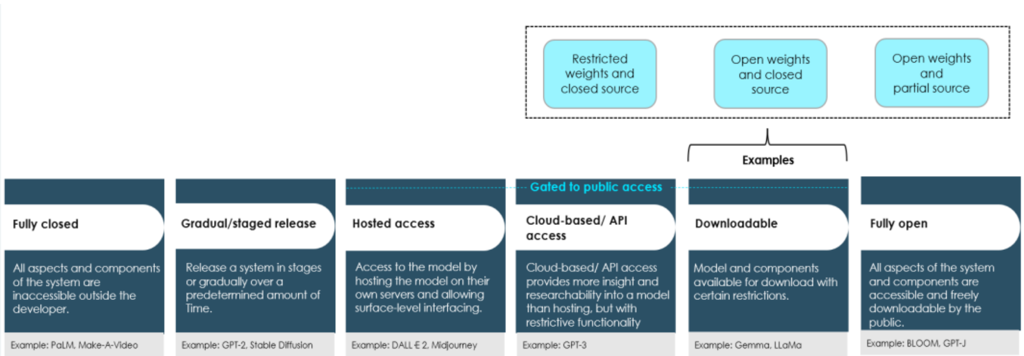
[11] appliedAI Institute for Europe (2023).
[12] Fine-tuning can allow more efficient performance or a specific focus on certain tasks or domains e.g. financial data, code-generation, or for applications dedicated to a specific client.
[13] For example, all the models fine-tuned based on the foundation model LLaMA released by Meta which gave access to the model's weights.
[14] The training code defines the model architecture and implements the algorithms used to optimise the model weights during training. The inference code implements the trained model, given the model weights and architecture; see Seger et al. (2023).
[15] Solaiman (2023).

12. There is no simple binary distinction between “closed” and “open” models but rather a spectrum of options from models that are accessible only through a developer’s systems to fully open models:

- Some firms develop their own proprietary AI models, which they use exclusively for their services without giving access to any aspects and/or components of the model outside of the organisation (e.g., BloombergGPT).
- Other firms provide access to their models via dedicated APIs, which allow third parties to query the model without access to the underlying code or detailed workings (e.g., OpenAI GPT-4).
- Developers may make available the model-trained weights, in other words, open-weight models (such as Google Gemma, Mistral Mistral7B and Meta LLaMA),[16] and/or parts of the code allowing other developers to build on or change the underlying model for their own needs.
- Finally, full-access models, such as BLOOM from Hugging Face, provide all the model information (including source code and the training datasets), thus allowing users to download all components and retrain the model if they wish.

13. Figure 2 presents an overview of the gradient of system access of GenAI models based on Solaiman (2023).

Figure 2
Degree of openness of GenAI models



Source: Copenhagen Economics, based on Salaiman (2023).

14. In this paper, we consider a foundation model “open” if at least its weights are publicly available, in line with the definition used by the United States (US) National Telecommunication and Information Administration in the public consultation on Dual Use Foundation Artificial Intelligence Models with Widely Available Model Weights.[17,18]

[16] Generative AI models generally estimate millions or billions of parameters that are used to determine the outcome of requests. Each of these parameters is given a different weight after the model is trained. Open-weight models share these weights.

[17] Council of the European Union (2024). Also see NTIA (2024).

[18] The draft version of the AI Act approved by the Council of Europe defines models as ensuring high levels of openness if *“their parameters, including the weights, the information on the model architecture, and the information on model usage are made publicly available. The licence should be considered free and opensource also when it allows users to run, copy, distribute, study, change and improve software and data, including models under the condition that the original provider of the model is credited, the identical or comparable terms of distribution are respected.”*

15. These models allow fine-tuning and application developers to freely build on them and develop new applications for end users. To date, the availability of open foundation models has been widespread within the industry, and this has played a significant role in driving progress in GenAI, allowing even firms with limited resources to access and develop products based on this technology (see Chapter 3).[19]

16. An assessment of GenAI needs to consider the added layers of complexity along the value chain. The same foundation model can be modified downstream in the value chain through fine-tuning or GenAI applications multiple times by its original developer or other developers in the case of open models. Different modified models can then be used by multiple entities further down the value chain (e.g. two fine-tuned models of the same foundation model can be used to develop various applications). It is also possible that a single GenAI application is based on multiple foundation models.

Competition authorities are keen to ensure that the GenAI race is competitive.

17. Several European competition authorities have launched general market studies of the still-forming GenAI market in a forward-looking approach to understand the market as it develops and ensure they can capture any competitive concerns quickly.

18. The Consumer and Markets Authority (CMA) was the first authority to publish its initial review of the role of foundation models in September 2023. The review, based on desk research and discussions with industry experts, assesses potential competition concerns and barriers to entry at the level of foundation models, as well as the possible impact of foundation models on competition in other markets. It includes several proposed principles to guide future competitive assessments to ensure markets are contestable and consumer interests are protected. However, the authority stressed that it is still too early to draw firm conclusions on how the market will develop.[20]

19. In addition to the CMA, the Portuguese Competition Authority (AdC) undertook an initial study of the GenAI market and published an Issues Paper on potential future concerns. The AdC did not find any existing concerns, but identified several areas where it considered that careful ongoing monitoring would be required.[21]

20. More recently, both the French[22] and Hungarian[23] Competition Authorities launched market studies, which they expect to complete during the course of 2024, while the Commission issued a call for contributions on competition in GenAI.[24]

21. Simultaneously, competition authorities have also launched behavioural investigations in markets considered crucial for the development of GenAI. For example, the French Competition Authority launched an investigation into Nvidia, the leading supplier of GPUs used to train the GenAI model, on suspicion of anticompetitive practices.[25]

22. Finally, some competition authorities have indicated the desire to assess agreements concluded between large digital market players and developers of GenAI. In particular, the European Commission,[26] the Bundeskartellamt[27] and the CMA[28] all announced potential merger investigations into the OpenAI -Microsoft partnership, see Table 1.

[19] Autoridade Concorrência (2023).
[20] The CMA (2023a) also indicated that they would publish a review of their assessment in 2024.
[21] Autoridade Concorrência (2023).
[22] French Competition Authority (2024).
[23] Hungarian Competition Authority (2024).
[24] For a review of competition issues in digital markets, see European Commission (2019) and European Commission (2024a).
[25] French Competition Authority (2023).
[26] European Commission (2024a).
[27] Bundeskartellamt (2023).
[28] CMA (2023b).

Table 1
European competition authorities' initial analysis of GenAI

AUTHORITY	INITIATIVE STATUS	DESCRIPTION
UK Consumer and Markets Authority (CMA)	• Issued initial report in September 2023. • New report forthcoming. • Potential merger investigation into the OpenAI-Microsoft partnership.	• Focus on understanding foundation models. • Too early to draw firm conclusions on how the market will develop. • Proposed principles for a competitive market. • Assessment review in 2024.
French Competition Authority	• Launched an initial investigation into the GenAI market in February 2024. • Launched an investigation against Nvidia.	• Gather stakeholders' views on practices implemented by major digital players that could harm the evolution of the GenAI market. • Initial stage: The outcome remains uncertain.
Portuguese Competition Authority (AdC)	• Issued paper on Competition and GenAI in November 2023.	• Map key determinants affecting competition. • Anticipate potential risks to competitive dynamics.
Hungarian Competition Authority (GVH)	• Launched market analysis on the impact of AI in January 2024.	• Focus on investigating the market behaviour of large technology companies and online platforms. • Initial stage: The outcome remains uncertain.
European Commission	• Called for comments on competition in GenAI in January 2024. • Potential merger investigation into the OpenAI-Microsoft partnership.	• Gather information and views in relation to competition aspects from industry, regulatory experts, academia, and consumer organisations. • Initial stage: The outcome remains uncertain.

Source: Copenhagen Economics desk research

2 POTENTIAL COMPETITION CONCERNS COULD ARISE FROM INPUTS, PARTNERSHIPS OR LEVERAGING

23. We draw on the economic literature on digital markets and the initial assessments undertaken by competition authorities to identify potential competition-related concerns in GenAI. In summary, the theoretical concerns identified are:

- The scarcity of key inputs may create barriers to entry in the development of foundation models.

- Partnerships between large firms and new players may stifle the emergence of new competitors.
- The conduct by large players may hamper competition in GenAI markets and other markets.

24. These theoretical risks serve as a framework within which we will assess current market developments in Chapter 3.

2.1 SCARCITY OF KEY INPUTS MAY CREATE BARRIERS TO ENTRY IN THE DEVELOPMENT OF FOUNDATION MODELS

25. Competition authorities have expressed concerns that access to key inputs – namely data, hardware, computing power, and technical expertise – may become a bottleneck for the development of new GenAI models, thus creating significant barriers to entry.[29]

26. Below, we set out each of the key inputs required, and where potential bottlenecks could arise due to lack of access.

27. **Training data:** Different types of data can be used at each stage of foundation model development, presenting implications for the risk that these may create barriers to entry.

    i. **Pre-training:** This is the stage requiring the most data. We understand that currently models are pre-trained mostly on publicly available data accessed via the web. , There are also several providers of datasets already pre-prepared for training. Moreover, some AI developers have already concluded licensing deals to access certain online proprietary content from popular websites such as Stack Overflow, Reddit, and news outlets. To the extent that new proprietary data may become important for the development of future models, firms that have access to large amounts of that data may become important suppliers.

    ii. **Fine-tuning:** This stage usually requires less, but often higher quality and domain-specific, data (e.g., financial data, conversations, medical records). Fine-tuning can be carried out by firms other than those that developed the foundation models and is more likely to use proprietary data. However, the smaller data requirements and domain-specific nature of these data suggest that the risk of creating barriers to entry is smaller. For example, startups offering AI tools to businesses may use the clients' data to create targeted fine-tuned models.

28. **Hardware:** The development of foundation models currently requires highly specialised hardware, namely GPUs. These GPUs are in short supply. Nvidia, the leading global supplier of GPUs arguably enjoys a near monopoly position. However, some cloud service providers (CPSs), such as Google and Microsoft, have also started developing their own chips and hardware to limit their dependence on Nvidia.[30]

29. **Computing power:** The development of foundation models requires large and sophisticated computing systems working for several weeks at a time. Access to these systems is currently in short supply and can be quite expensive. Pre-training of these models could cost several million dollars (the cost for training GPT-4 was reported above $100 million, while Meta estimated a cost of $4 million for LLaMA).[31]

---

[29] Significant barriers to entry could imply that the market for foundation models (or of GenAI applications dependent on them) becomes concentrated. We note that other characteristics such as economies of scope or the presence of network effects may also lead to concentrated markets.

[30] See for example *The Verge* (2023).

[31] We note that these numbers are not necessarily comparable and may not include additional costs linked to the development of those models. See *Forbes* (2023) and Medium (2023a).

30. Most AI developers do not have their own data centres and instead rent computing power from CPSs such as Microsoft Azure, Google Cloud, AWS, and a few small players. Some GenAI developers may enter into long-term agreements and/or complex partnerships with CPSs to secure continuous access to computing power at reduced prices (see Section 2.2).[32] These partnerships may be required for developers to guarantee that they have access to this important input at an affordable price, while also accessing funds and limiting the risks that access to these inputs may be lost in the long term.

31. **Technical expertise**: Due to their complexity, the development of GenAI models requires a high level of technical expertise. The technical expertise required includes cutting-edge knowledge of machine learning as well as significant practical expertise in data engineering and high-performance computing.

32. **Funding:** Finally, developing a foundation model requires significant funds in order to gain access to the inputs identified above.

2.2 PARTNERSHIPS BETWEEN LARGE FIRMS AND NEW PLAYERS MAY STIFLE POTENTIAL COMPETITION

33. Partnerships between existing large tech firms with access to key inputs, particularly cloud services, and startups developing AI foundation models are common. For example, Microsoft and OpenAI entered into a partnership in 2019, while Anthropic entered into a partnership with both Google and AWS in 2023.

34. The main rationale for these agreements is for the AI startup to gain access to the computing power required to train its model(s), as well as external funding and distribution channels (e.g., the cloud platform operated by the cloud provider).

35. The nature and depth of these agreements vary, and they may range from simple agreements to share infrastructure to closer partnerships, which can include sharing of knowledge, privileged access for the larger tech company to the startup's foundation models (such as, for example, exclusive IP licensing rights or exclusive rights for distribution of certain top-tier versions of its models), or even some level of control (which may be decisive) over the startup.

36. Because the largest firm in the partnership often has access to a key input and distribution channel, it may obtain favourable terms, which could create competition concerns. The competition concerns that may arise from these partnerships depend on a variety of factors, including the relative position of both competitors, the level of control, and the terms and conditions of the agreement. One potentially important concern is where agreements may remove potential competition from the startup itself in markets where the large tech firm may be present or dominant.

2.3 LEVERAGING BEHAVIOURS BY LARGE FIRMS MAY HAMPER COMPETITION IN GEN AI MARKETS AND OTHER MARKETS

37. Standard competition law has grappled with leveraging behaviours in digital markets and beyond.[33] Leveraging exclusionary conduct occurs when a firm with market power in one market uses it to exclude potential rivals in a related market, thus gaining an unfair competitive advantage. These conducts can take many forms, such as self-preferencing, refusal to supply, or tying/bundling.[34]

---

[32] These partnerships may be required for the developers to guarantee that they have access to this important input at an affordable price, gain access to funding, and limit the risks that access to these inputs may be lost in the long term.

[33] See past cases such as Google Shopping (i.e., case AT.39740) also discussed by Motta (2023), Google Android (i.e., case AT.40099), Amazon Buy Box (i.e., case AT.40703), and Apple Pay (i.e., case AT.40452).

[34] See Motta (2023) and Motta and Fumagalli (2024).

38. While the exact form of anticompetitive conduct differs from case to case, the common thread is that the market where there is market power is closely related to the affected market (in this case, a Generative AI market).[35] This can be because it provides an important input, key customers, or it is a product that is typically bought together or can be integrated. The leveraging behaviour makes it harder for rivals in the affected market to access the input, customers, or integration, and this, in turn, gives the integrated firm market power in that market.[36]

39. In general, the effect of these conducts must be assessed on a case-by-case basis, starting from the important question of whether the firm has market power in the market used for leverage. The assessment should also consider pro-competitive benefits that may arise from coordination and integration, such as better integration of products, improved customisation, etc.

40. In markets and practices not covered by the DMA, competition authorities may have to rely on behavioural competition law, such as the prohibition covered under Article 102 of the Treaty on the Functioning of the European Union (Article 102 TFEU). Authorities may consider using measures that allow them to intervene quickly, such as settlements or interim injunctions.

41. Recent *ex-ante* regulations such as the Digital Market Act (DMA) may capture some of the potential leveraging conducts if a) the firm adopting the conduct is a designated gatekeeper and b) the service that integrates GenAI features is recognised as a core platform service (e.g., search engines, operating systems, virtual assistants, etc.).

3 GENAI MARKETS ARE VIBRANT, WHILE FORECLOSURE IS STILL A RISK

43. In this chapter, we assess whether and to what extent recent market developments allow us to identify any of the potential competition risks discussed in Section 2.

44. First, we explain that the high number of active players in the GenAI market and the speed at which new models are currently launched suggests that there is no sign that significant barriers to entry exist. The lack of signs of market power at the level of foundation models indicates that the risk of anticompetitive leveraging from foundational models to downstream or adjacent markets is also limited.

45. Second, we review partnerships between large integrated players and AI startups. We explore the intricacies of existing partnerships based on publicly available information on the contracts.

46. Finally, in the third section, we assess the integration of GenAI applications into existing services from large digital firms and the potential competitive impacts of these decisions. We explain that while non-integrated new players remain competitive, this leveraging behaviour should be monitored closely.

3.1 DEVELOPMENTS IN GENAI SUGGEST BARRIERS TO ENTRY ARE NOT CURRENTLY AN IMPEDIMENT

47. In this section, we present our findings on the current high-level indicators of the competitive functioning of markets. We find the following: There is a large and increasing number of firms developing foundation models. The field includes both new startups and established firms in other digital markets. Additionally, several solutions with different degrees of openness are available,

---

[35] Note that if market power exists in foundational models, it could be the case that both the market used for leveraging and the leveraged market are markets related to GenAI.

[36] See Katz (2018).

which means that firms wishing to develop new downstream applications have easy access to foundation models.

48. Additionally, there is a large inflow of venture capital investments to the GenAI sector to fund new enterprises, while technical advances are making training and using GenAI models more computationally efficient and thus cheaper.

49. Finally, we highlight a set of uncertainties that may affect future developments in the GenAI markets.

50. The evidence presented in this section is not consistent with the presence of strong barriers to entry, or firms with access to key inputs having a significant competitive advantage. There are no signs that this market will tip to only one or a limited number of companies. In turn, this also reduces the risk of anticompetitive leveraging behaviours of foundation model developers at the deployment stage and/or in adjacent markets.

3.1.2 There is a large and increasing number of foundation model developers

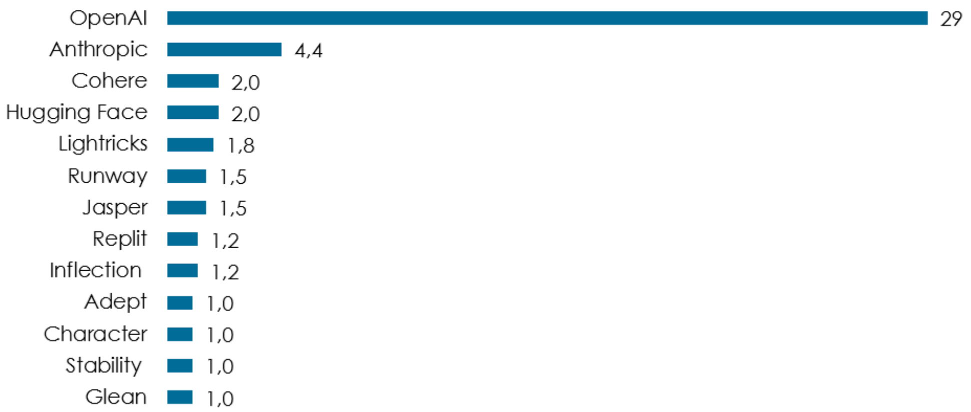
51. According to a Standford database,[37] more than 250 foundation models have been developed since 2018 by 94 different organisations, and 57% are available on an open licence. Using the same source, the CMA reported 160 models in August 2023, meaning that in only six months since then, more than 100 new models were introduced.[38]

52. New successful startups have emerged in the GenAI sector, chief among them OpenAI (see the case study in Box 2 below). As of May 2023, there were already 13 startups that obtained a $1 billion valuation (also named “unicorns”) in the GenAI sector according to private equity firm CB Insights, see Figure **3**. The pace of development is so fast that new unicorns have already joined the field since, such as the French firm Mistral.[39]

53. Many of these startups have developed their own foundation models. Some of them have developed all-purpose conversational chatbots similar to OpenAI’s ChatGPT (e.g., Anthropic’s Claude), while others specialise in video creation (Runway), coding (Replit), automated actions (Adept), or AI applications and communities (Hugging Face).

[37] See Stanford University (2024a). We note that this may also include fine-tuned models and different versions of the same model. The database was accessed in February 2024.
[38] We note that there seems to be a slight discrepancy of around 20 between the number reported by the CMA in its publication in September and our calculations for the models reported up to CMA access.
[39] *The New York Times* (2023).

Figure 3
AI startups with $1 billion+ valuations as of May 2023



Source: Copenhagen Economics based on CB Insights (2023). We note that since May 2023, new startups such as Mistral[40] have reached $1+ billion valuation, and the valuation of some of these firms has already increased, e.g. OpenAI is now valued at around $86 billion.

54. European players are also becoming increasingly important in the sector, including leading players such as Mistral (see Box 3), and Aleph Alpha (see Box 4). According to the appliedAI Institute for Europe,[41] there are approximately 6,300 AI startups in the EU, of which approximately 10.6% can be classified as GenAI startups. Some of these GenAI startups develop their own foundation models or downstream applications, while others provide development tools and critical infrastructure for GenAI purposes.[42]

55. Figure 4 lists the foundation models for different applications developed by a large number of firms. The colour coding indicates different levels of model openness (as discussed in Section 1): fully open models with access to source code and weights, open models with access to weights, closed models accessible only via an API, or fully closed models.[43]

[40] *The New York Times* (2023).
[41] appliedAI Institute for Europe (2023).
[42] These GenAI Startups are located across the EU, with Germany (19.9%), France (17.5%), the Netherlands (10.9%), and Sweden (8.2%) leading the way.
[43] Open models are freely available for further specialisation by any firm, in particular, the code which trains the model, the trained weights for fine-tuning or application development. Closed source, available through APIs, means that the model's source code and the trained model are not publicly available ,but the developer allows other users to connect to the model through an application programming interface that allows the developers to use the model's features for their own programs. Finally, 'closed source' means that the foundation model is only available to its original developer.

Figure 4
Example of GenAI models

	Text	Image	Audio/music	3-D	Video	Protein structures or DNA sequences
Microsoft			VALL-E	RODIN Diffusion	GODIVA	MoLeR
OpenAI	GPT-4	DALL-E 3	Jukebox	Point-E	Sora	
Meta	LlaMA 2	Make-a-scene	AudioGen	Builder Bot	Make-a-video	ESMFold
Google	Gemma 2B/Gemma 7B Gemini 1.5	Gemini 1.5	Gemini 1.5	Gemini 1.5	Gemini 1.5	Isomorphic labs
Amazon	Titan	Titan	DeepComposer			
Stability AI	StableLM	Stable Diffusion 2	Dance Diffusion		Stable Video Diffusion	LibreFold
Nvidia	MT-NLG	Edify		Edify	Edify	MegaMolBART
Midjourney		Midjourney V.6				
Cohere	Family of LLMs					
Anthropic	Claude 2					
AI21	Jurassic-2					
Mistral AI	Mixtral 8x7B Mistral Small Mistral Large					
Aleph Alpha	Luminous	Luminous				
01.AI	Yi-34B					
Runway					GEN-2	

Closed source | Closed source, available through APIs | Open weight | Open source

Source: Copenhagen Economics based in part on McKinsey & Company (2023). The list is not exhaustive.

56. Models from new startups now compete head-to-head with models developed by established players despite the lack of access to the same inputs or funds. The platform LMSYS (developed by Hugging Face) evaluates large language models according to a set of benchmarks. According to its latest results, the best-performing model is currently OpenAI GPT-4, followed by Google's Gemini models, Mistral's Medium model, and Anthropic's Claude models.[44]

57. Finally, open models remain an important alternative for AI developers who wish to develop fine-tuned models for specific tasks and domains. These models therefore place a competitive constraint on all downstream applications and fine-tuned models developed by the foundation model developers. This reduces the risk that the market will tip towards a limited number of non-publicly available models. Recent developments suggest that open models have managed to achieve levels of performance comparable to large, popular closed-source models.

58. The importance of open foundation models such as Gemma (Google), Mistral 7b (Mistral), and LLaMA (Meta) is mentioned by the US National Telecommunication and Information

[44] LMSYS Org (2023).

Administration in its public consultation on Dual Use Foundation Artificial Intelligence Models with Widely Available Model Weights:

> "[…] foundation models with widely available weights […] could play a key role in fostering growth among less resourced actors, helping to widely share access to AI's benefits. Small businesses, academic institutions, underfunded entrepreneurs, and even legacy businesses have used these models to further innovate, advance scientific knowledge, and gain potential competitive advantages in the marketplace. The concentration of access to foundation models into a small subset of organizations poses the risk of hindering such innovation and advancements, a concern that could be lessened by [the] availability of open foundation models. Open foundation models can be readily adapted and fine-tuned to specific tasks and possibly make it easier for system developers to scrutinize the role foundation models play in larger AI systems, which is important for rights- and safety-impacting AI systems (e.g., healthcare, education, housing, criminal justice, online platforms etc.)"[45]

59. Box 2 discusses the American firm OpenAI, and Box 3 and Box 4 discuss the European firms Mistral and Aleph Alpha, three firms that have shaped the developments of the GenAI space.

[45] See NTIA (2024).

Box 2 Case Study: OpenAI's success story

OpenAI was founded in December 2015 by Elon Musk, Sam Altman and other investors as a non-profit AI research lab, with the stated aim of advancing GenAI for the benefit of humanity. It later transitioned to a "capped-profit" model. In 2019, OpenAI received a first initial investment of $1 billion from Microsoft. In November 2022, OpenAI publicly launched the conversational chatbot ChatGPT, which generates natural and coherent responses to user inputs. ChatGPT became an instant phenomenon, setting a record for having the fastest-growing user base in history for a consumer application (100 million monthly active users in just two months).

ChatGPT is based on GPTs, a set of proprietary foundation models designed with transformer architecture (initially pioneered by Google researchers). OpenAI has also developed other models, such as the image generation model DALL-E 2.

In January 2023, OpenAI received another $10 billion in investments from Microsoft, which became its major investor and only supplier of cloud services. Additionally, Microsoft can now use OpenAI's technology with a level of exclusivity.

In March 2023, OpenAI released GPT-4 with human-like capabilities, such as answering reading and maths questions, beating all benchmarks for AI models.[46]

In early 2024, OpenAI launched the GPT store, an online marketplace for custom chatbots derived from ChatGPT.[47]

OpenAI was recently valued at $86 billion[48] with an annualised revenue in 2023 exceeding $2 billion, driven by the widespread adoption of its premium ChatGPT product.[49] OpenAI stands among the few Silicon Valley companies (including Meta and Google) that have reached $1 billion in revenues within a decade of being founded.

The Open AI business model is centred around two main products: ChatGPT, and its foundation model (GPT-4). Customers can use the ChatGPT free version or pay a subscription (around $25-30 a month per user) to have access to more functionalities and better performance with multiple options for businesses – ChatGPT Enterprise[50] and ChatGPT Team.[51] OpenAI also gives access to its GPT models to other developers via dedicated APIs on which it can charge a fee.

OpenAI has gained an early lead over large digital incumbents (such as Google, Amazon, and Meta), and its models are still leading across most available benchmarks. In particular, with the launch of ChatGPT, OpenAI became the leader in GenAI chatbots and established its consumer brand. It has different types of deployment and monetisation strategies, from open source to open access via API (i.e., no disclosure of model specifications, training datasets, etc.). It is now rumoured that OpenAI plans to integrate a search engine functionality to directly challenge Google, and it plans to expand into computer assistants capable of executing varied actions.

---

[46] OpenAI (2023a).
[47] OpenAI (2024a).
[48] *Financial Times* (2023a).
[49] *Financial Times* (2024a).
[50] OpenAI (2023b).
[51] OpenAI (2023c).

Box 3 Case study: European Mistral offering open-source competitive solutions

Founded in 2023 by European former employees of Google's DeepMind and Meta, the French startup Mistral is committed to serving the open community and their enterprise customers by providing open-weight generative models that match the quality of proprietary models. Mistral had a meteoric rise: it succeeded in raising $113 million in seed funding within just four weeks from inception, and after another round of funding of $500 million in December 2023, is now valued at around $2 billion.

This success is due to its extremely good models. Despite lacking proprietary access/advantage in hardware and/or data, Mistral managed to compete effectively with the much larger models of some competitors (such as GPT-4), and now Mistral models are considered by the AI community among the best-performing ones.

On 26 February 2024. Mistral entered into a multi-year partnership with Microsoft. This partnership provides Mistral AI with access to Azure's AI infrastructure (i.e., Microsoft's cloud computing service).[52]

Mistral produces models with different degrees of openness: from open-source models to models provided via API:

- Mistral-7B (September 2023): an advanced open-source language model. This community-backed model allows us to quickly add capabilities (e.g., context life extension, image encoders, direct preference optimisation, etc.). The model outperformed other renowned models (such as LLaMA 2) across a range of benchmarks, especially in tasks involving reasoning, mathematics, and code generation.[53] With just 7 billion parameters, Mistral-7B delivers top-tier performance at a lower cost than models with more parameters (i.e., it requires nearly 50% less computational power to run).[54]
- In December 2023, Mistral introduced a new mixture of expert models (MoE), Mixtral 8x7B. It stands as the most robust open-weight model with a permissive licence and holds the top position overall in terms of cost/performance trade-offs. Specifically, it either matches or surpasses GPT 3.5 across the majority of standard benchmarks.[55]
- In February 2024, Mistral introduced its most advanced LLM to date, Mistral Large. The model is available through Mistral's own API through La Plateforme and Microsoft Azure. The model achieves strong results across a range of benchmarks, making it the world's second-ranked model generally available through an API.[56] At this time, Mistral also announced its new chatbot, Le Chat.[57]
- At the same time, Mistral announced a new strategic partnership with Microsoft, including limited investment and a distribution agreement via Microsoft's CPS Azure.[58]
- Mistral relies on various monetisation strategies. On the one hand, its open source enables the company to leverage improvements and iterations on the model contributed by the AI members and the machine learning community. On the other hand, Mistral develops proprietary models tailored to specific business clients. Currently, Mistral is directing its focus towards applications such as summarising company communication, document querying, and generating personalised marketing material.[59] It recently entered into a partnership with Capgemini to expand the adoption of its models.[60]

---

[52] *Financial Times* (2024b).

[53] Mistral AI (2023a).

Box 4 Case study: Aleph Alpha is an emerging EU player

Aleph Alpha, a German startup founded in 2019, aims to create a European alternative to OpenAI, and thus become the foremost European firm in the AI space. Aleph Alpha prides itself on its data protection, security, and transparency credentials. The startup secured $500 million in Series B funding in November 2023, representing the seventh-biggest deal for all AI companies in 2023.[61]

The startup allows companies – particularly in sensitive and regulated industries–- to develop and deploy their own large language and multimodal models:

- In 2022, Aleph Alpha developed Luminous, a family of multimodal language models. It includes three models, each varying in complexity and capability. According to Aleph Alpha, Luminous is on par with some of the world's leading AI language models, while being significantly more efficient.[62]
- In June 2023, Aleph Alpha launched its new generation of Control-Models. For every Luminous model, there is now a control version available that can be used by its partners and customers. The models are designed to enhance natural language processing and tackle computational linguistic tasks.
- In August 2023, Aleph Alpha developed, together with Bosh (a leading global technology and services supplier), BoshGPT, an AI language model similar to ChatGPT.[63]

3.1.3 The high level of investments in the Generative AI space suggests that investors trust its market potential

60. GenAI is the fastest-growing segment of AI and has attracted corporate and venture capital (VC) funding globally.[64] This influx of investments reduces barriers to entry as it allows firms to acquire key inputs. Moreover, the willingness of the venture capital sector to invest money in a significant number of new players suggests that it believes in the sector's potential.

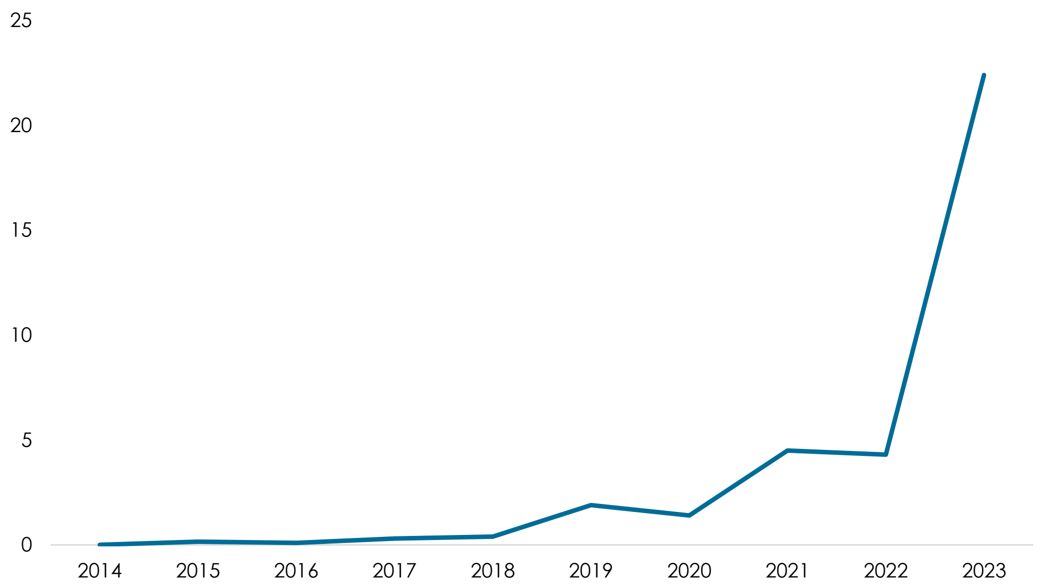
61. In 2023, the AI sector raised $50 billion in venture capital globally, despite a broader slowdown in venture investments.[65] This represented almost a quarter of all corporate VC investments. [66] According to Goldman Sachs, investments in the AI sector are projected to reach $200 billion globally by 2025.[67]

---

[54] *Sifted* (2023a).
[55] Mistral AI (2023b).
[56] Mistral AI (2024a).
[57] Mistral AI (2024b).
[58] *Financial Times* (2024b).
[59] *Sifted* (2023a).
[60] Capgemini (2024).
[61] Out of a total of $500 million, $120 million consists of equity. See CNBC (2023) and Dealroom (2023).
[62] Aleph Alpha (2023).
[63] TheNorthAI (2023).
[64] We note that a significant share of investments comes from large tech players such as Microsoft, Google, and Amazon. According to the *Financial Times* reporting, investments from Microsoft, Amazon, Nvidia and Google (Mang) were about 30 per cent of the total investments in data and AI start-ups. The share may be higher if looking at only large deals in fledgling GenAI companies. See *Financial Times* (2023b).
[65] Air Street Capital & Benaich (2023).
[66] Air Street Capital & Benaich (2023).
[67] Goldman Sachs (2023b).

62. GenAI now accounts for almost half of the total investment in AI in 2023.[68] Investments in GenAI reached $22.4 billion in 2023, up from only $4.3 billion in 2022, marking an increase of 422% (see Figure 5).

Figure 5
Worldwide venture capital investments in GenAI
$ billion



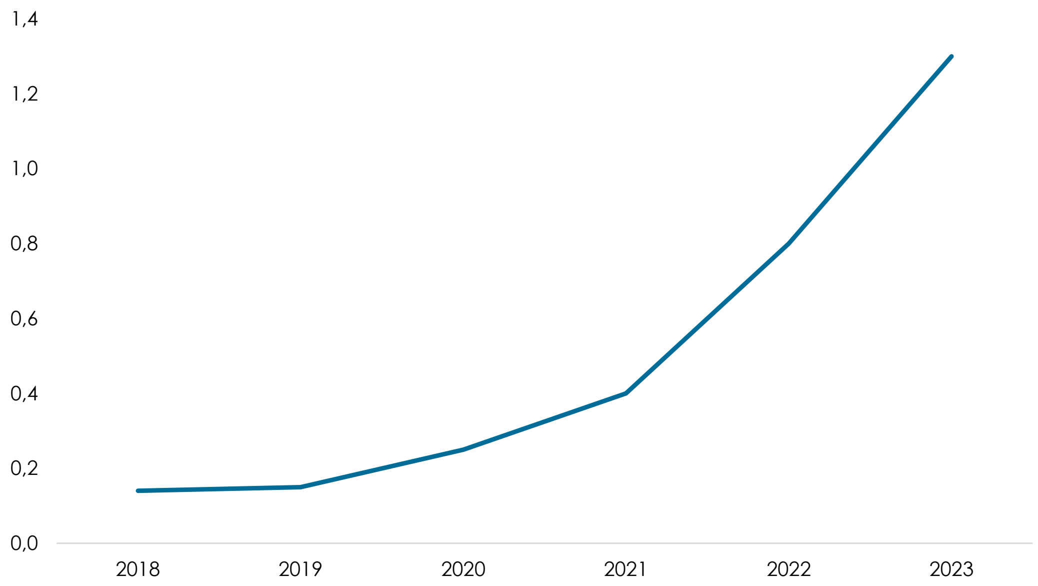
Source: Datawrapper (2024).

63. In Europe, AI startups (which include GenAI) saw a 60% increase in funding in 2023 compared to 2022, primarily fuelled by large-scale investment rounds such as Mistral's $113 million seed round and Aleph Alpha's $120 million Series B round. In 2023, venture capital investment in European Novel AI startups amounted to $1.3 billion (see Figure 6).[69] Another source, the appliedAI Institute for Europe, reported that European GenAI startups have so far received approximately $2.57 billion in overall private funding.[70]

64. Europe is also catching up on AI research, which can be a precursor to startup formation and new investment VC opportunities. According to (Maslej, et al., 2023) between 2010 and 2021 institutions in the European Union and the United Kingdom accounted for 20% and 15% of the global AI conference and journal publications, respectively.

[68] Air Street Capital & Benaich (2023).
[69] Dealroom (2023). It is unclear whether this figure already includes the $400 million secured by Mistral in its second round of investment in December 2023.
[70] appliedAI Institute for Europe (2023).

Figure 6
VC investment in European novel AI startups over time
$ billion



Source: Dealroom (2023).

65. In addition to private investments, governments around the world have shown an interest in supporting this industry. A global race for the development of national alternatives has started, with different types of subsidies offered to firms that develop their GenAI products in each country.[71]

3.1.4 Computing demand is expected to decrease as the focus shifts to fine-tuning and more efficient models

66. GenAI models often require vast amounts of expensive and scarce computing power. In recent years, important technological advances that enhance the cost-effectiveness of models have emerged, particularly in the realm of fine-tuning. According to some commentators, this will allow smaller companies with no access to computing power to compete with Big Tech giants.[72]

67. There are indications that models will not continue to grow indefinitely. OpenAI's CEO Sam Altman recently said in an interview that the age of giant AI models is over, suggesting that developers will from now on "make [them] better in other ways".[73] For example, the new Large model developed by Mistral performs almost on par with the older OpenAI GPT-4, while reportedly substantially "smaller" in its model weights.[74]

[71] A recent article for The Economist states that *"[t]he proliferation of LLM-makers reflects a belief that the market for AI will be geographically fragmented and hypercompetitive." The Economist* (2024a). Similarly, an article in the FT by a special adviser to the European Commission states that *"[o]n both sides of the Atlantic, feverish public investments are being made in an attempt to level the computational playing field." Financial Times* (2024c).

[72] Schrepel & Pentland (2023).

[73] *Wired* (2023b).

[74] See *The Economist* (2024b).

68. There are also several recent developments in GenAI models, which suggest that the computing power required to develop comparable new models is decreasing significantly, see Box 5.

Box 5 Example of promising innovations to increase computational efficiency

- The low-rank adaptation (LoRa) is a training technique that relies on the low-rank decomposition of weight matrices. This approach enables a significant reduction in the number of trainable parameters (by 10,000 times) and GPU memory requirements (by three times).[1]

- Step-by-step distilling is a technique for fine-tuning smaller language models. It involves extracting informative natural language rationales from LLMs, which can, in turn, be used to train small task-specific models. This mechanism allows the training dataset to be reduced and leads to smaller models outperforming few-shot prompted large language models (LLMs) with 700 times more parameters.[2]

- The Small Language Model (SML) explored by, among others, Microsoft aims to achieve capabilities similar to LLMs, such as OpenAI's GPT-4, but with reduced computing power requirements. The company has already demonstrated its ability to compete with 50 times larger models.[3]

- Gemini models, released by Google,[4] offer different data and computing capacities across various model sizes (ranging from Gemini Pro to Gemini Mini). These models will most likely continue to co-exist, but with different purposes.[5]

- The 'Mixture of Experts' (for instance, Mixtral developed by Mistral[6]) is a technique in which a query triggers only certain portions of the model, thus making the model computationally efficient and performing better than other, larger, models.[7]

Source: [1]Shen et al. (2021), [2] Hsieh et al. (2023), [3]*The Information* (2024a), [4] Google (2024b), [5]Engadget (2024), [6]Mixtral (2023b), [7]Medium (2023b).

69. Finally, policy initiatives have emerged to support developments in this space. For example, the recent EU's AI innovation package[75] will support SMEs and startups in accessing some of the key inputs identified above. SMEs that qualify for this model will gain access to EU supercomputers and high-quality datasets to train their models.

3.1.5 Despite a promising start, uncertainties remain on the potential future for innovative firms

70. The evidence shown in the remainder of this section suggests that barriers to entry are not, currently, a significant impediment to the development of new foundation models and/or GenAI applications.

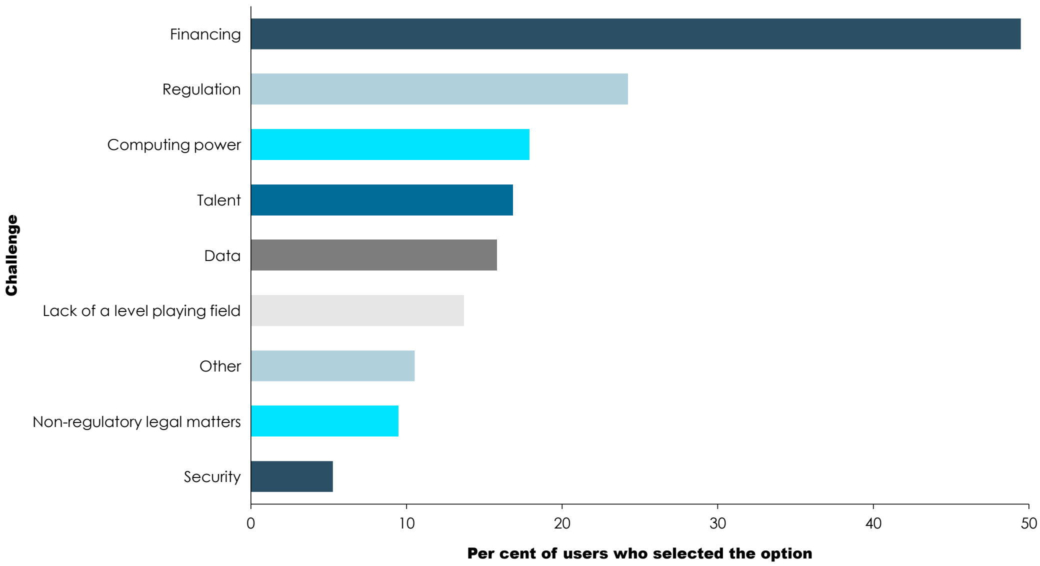
71. However, uncertainties remain as to future developments at this stage. For example, according to research by the appliedAI Institute, European startups still face significant challenges in entering the GenAI market. The main challenges identified related to access to financing, as well as the burden brought by excessive regulatory interventions, see Figure 7. Given this, particular

[75] European Commission (2024b).

care should be given to the impact of any future regulations on the development of Generative AI, particularly on potential impacts on costs for smaller firms who are typically less able to bear any regulatory costs. European AI startups were specifically mentioned in the discussions surrounding the recent EU AI Act due to the potential burden that compliance with this regulation would entail.[76]

72. Future regulatory activity in the area of copyright and data privacy may also affect AI startups' ability to collect and process publicly available data. This debate has already given rise to litigation in the US against developers of large foundation models for their use of publicly available data.[77] Any changes in the ability to use publicly available data will likely have a larger impact on startups with limited resources.[78]

Figure 7
Major challenges of European GenAI startups
Number of respondents who chose each option



Note: The 95 respondents present in the sample were allowed to choose multiple options. In the original report, the third category is "compute power" which is typically used in developer documents to refer to "computing power". The source data for the graphic is not publicly available and therefore the percentages are estimated based on the chart in [source].
Source: appliedAI Institute for Europe (2023).

73. In addition, more than 10% of European startups surveyed identified access to computing power, talent, and data as a potential challenge. Given this, the potential impact of these inputs may be still not fully understood. We consider some of these uncertainties below.

74. As explained above, the computing power required for developing models is decreasing. However, the demand for increasingly complex models may increase faster. The relative speed at which these two factors evolve will determine whether the total computational cost of new foundation models increases or decreases.

[76] *Sifted* (2023b).
[77] See MLex (2024) and Stanford University (2024b).
[78] We understand that currently AI models are able to access and use publicly available data with copyright through the European Copyright's Directive Text and Data Mining exception. This exception gives rightsholders the ability to opt out their content from this exception when the product is used for commercial ends.

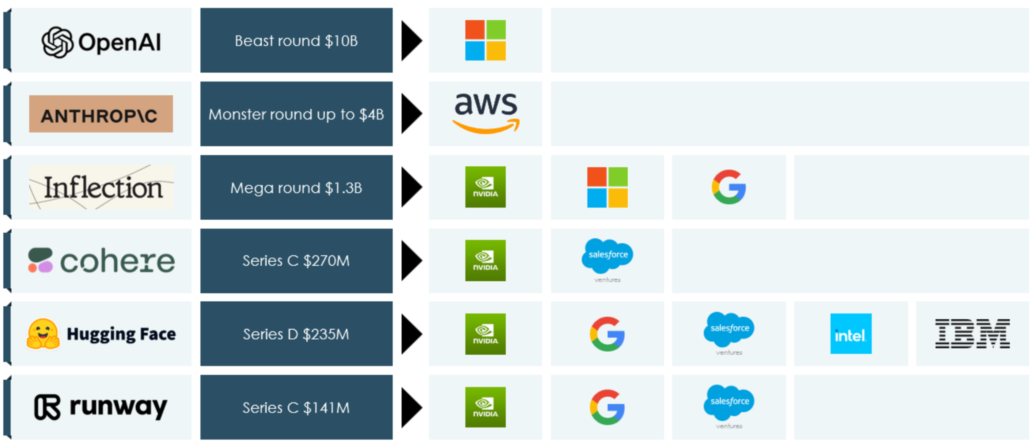
75. Finally, the role of proprietary data in future GenAI models is not yet fully understood. While current top-performing models have been developed that rely almost exclusively on publicly available data, the potential to train better models exclusively on public data may be exhausted in the future. At this stage, it is difficult to predict whether and what proprietary data may become most valuable for the future development of models. It is also important to note that different firms may be able to compete effectively by using alternative sets of proprietary data. The impact of new legislation designed to increase access to data (such as the Data Act, the DMA or the Data Governance Act) should be taken into account when assessing the need for further intervention.

3.2 PARTNERSHIPS PRESENT DIFFERENT DEGREES OF INTEGRATION AND MUST BE ASSESSED ON A CASE-BY-CASE BASIS

76. As discussed in Section 2.2, several smaller startups have entered into broad partnership agreements with large digital players, particularly those with access to key cloud computing power, in order to improve their chances of success.

77. Companies such as Microsoft, AWS, Nvidia, Google, and Salesforce have been present in the major funding rounds of the most prominent AI startups, see Figure 8.

Figure 8
Some of the highest-profile AI fundraises in 2023



Source: Air Street Capital & Benaich (2023). We note that Google also invested in Anthropic in 2023 (it has committed around $2bn in Anthropic according to (Reuters, 2023b) and Microsoft invested in Mistral in February 2024.

78. The race to support innovative firms in the GenAI value chain is not unique to large digital players, and, naturally, these firms may see added value in forming strategic partnerships with suppliers in markets that are very close to their main business activities.

These partnerships give GenAI developers unimpeded access to important cloud infrastructure that they require to develop their foundation models and GenAI tools, while giving large digital players access to tools that they can use in their own services (e.g., both Google and Amazon offer foundation models from multiple developers through their CPS).

79. An important consideration will be whether these agreements may dampen potential competition at the foundation model or downstream levels because of their restrictive conditions, or because they mimic a merger situation by granting the large digital player significant (or decisive) control over the startup. That control may, for example, impede the startup from competing in markets where the large digital player is present, or affect its incentives to supply its tools to the large digital players' competitors.

80. For example, certain news reports suggest that Microsoft/OpenAI gives Microsoft the exclusive rights to provide cloud computing services to OpenAI, as well as certain exclusive rights over OpenAI's intellectual property.[79] To the extent that this intellectual property becomes an important input for potential rivals to develop competing GenAI applications, this may result in a reduction in competition. This agreement is also being reviewed by some competition authorities due to the level of control that it may grant Microsoft over OpenAI.

81. Each partnership is unique and needs to be assessed individually. A review of publicly available news reports suggests that Microsoft's agreement with OpenAI is more closed than the agreements Anthropic has signed with both Google and Amazon, see Table 2.

82. On 26 February 2024, Microsoft announced a partnership agreement with the French startup, Mistral. Based on initial reports, this agreement does not have any exclusive rights for Microsoft and does not grant Microsoft control over Mistral.[80]

[79] See *Financial Times* (2024b).

[80] See *Tech Crunch* (2024).

Table 2
High-level comparison of OpenAI and Anthropic partnerships with large digital players

PARTNERSHIP CHARACTERISTICS	OPENAI – MICROSOFT	GOOGLE - ANTHROPIC	AWS - ANTHROPIC
Size of investment	Multiple investments between 2019 and 2023 reported around $13 bn in total.	Google is reported to have invested around $2 bn.	AWS is reported to invest up to $4 bn.
Control	Microsoft can appoint a board member of OpenAI, although only in an observing capacity.	Google is reported to have only 10% ownership in Anthropic.	AWS is reported to have minority ownership in Anthropic.
Profit sharing	Microsoft allegedly gets a significant proportion of OpenAI's profits for the foreseeable future to repay the investment. Microsoft and OpenAI allegedly receive co-payments from sales of OpenAI via the Azure platform and from sales of Microsoft GenAI services powered by OpenAI models.	Google is reported to have 10% ownership in Anthropic. Google allegedly received a share from the sales of Anthropic's models made through their cloud platforms.	AWS allegedly received a share from the sales of Anthropic's models made through their cloud platforms.
Choice of cloud provider	Exclusive supply from Microsoft -Azure.	Non-exclusive supply from Google (and potentially other providers).	Non-exclusive supply from AWS (and potentially other providers).
Choice of model distribution	The GPT model is only offered to Azure customers and OpenAI's direct sales channel. It cannot be sold through other CSPs.	The Claude model is offered on Google Cloud non-exclusively.	The Claude model is offered on AWS non-exclusively.
Access to model technology	Microsoft has an exclusive agreement to integrate OpenAI models in its services (e.g., Copilot powered by GTP-4). The degree of Microsoft's access to OpenAI technology is not clear.	Google likely has no access to Anthropic technology.	AWS likely has no access to Anthropic technology.

Note: The details of the agreements remain confidential. The information provided in the table is based on media reporting.
Source: Copenhagen Economics based on desk research and news articles. See *The Information* (2024b), Amazon (2023), *Financial Times* (2023c), *Reuters* (2023b).

3.3 DEPLOYMENT IS STILL AT AN EARLY STAGE WITH SOME LARGE PLAYERS INTEGRATING GEN AI INTO THEIR EXISTING SERVICES

83. Downstream deployment of new GenAI applications by businesses and final users is still at an early stage. In particular, the latest figures suggest that GenAI has not yet achieved the level of widespread adoption that experts believe it will be able to achieve. Currently, there is little or no sign of a risk of leveraging market power at the level of the foundational AI model.

84. First, most foundation model developers (e.g., OpenAI, Anthropic, Mistral) already provide their own user-facing applications. At the same time, many foundation model developers offer rival developers the ability to build applications on top of their existing models via open models or closed APIs (e.g., Jasper.ai writing assistant built on OpenAI GPT models).

85. Second, players in different levels of the value chain often specialise in specific domains (e.g., BloombergGPT in finance, or Isomorphic Labs in the pharmaceutical drugs discovery process) or tasks (GitHub coding copilot, or Perplexity.ai in search). The success of these models suggests that specialisation is a viable business model.

86. GenAI applications can provide an opportunity for challengers in already existing markets to compete against the more established players and increase the level of innovation in these markets. For example, Microsoft has integrated its GenAI output into its Bing search engine in a move to compete with Google. This also creates, however, a risk that players active in important adjacent markets to the GenAI market may be able to leverage that market power to GenAI applications, as discussed in Section 2.3. GenAI tools can be used in several domains, where some firms currently may enjoy a certain degree of market power.

87. Several firms have already integrated GenAI applications into their products. For example, both Microsoft and Google now offer integrated versions of their business productivity suites with GenAI models. Equally, Google, Microsoft, and AWS offer access to their and their partners' foundation models through their cloud services.

88. When GenAI applications are provided by a designated gatekeeper and integrated into core platform services (such as operating systems, search engines, or virtual assistants), they will be subject to the rules of the DMA. They will also be subject to these rules if they are integrated into new core platform services and meet the thresholds to be designated as a gatekeeper (e.g. if OpenAI were to launch its own search engine and this engine met the DMA's thresholds).

89. Outside of the scope of the DMA, leveraging conducts can also be assessed under Article 102 TFEU or its national equivalents. This assessment requires a careful case-by-case assessment to ensure they do not result in foreclosure. This effect-based analysis should consider the potential for impact on the relevant markets and the presence of dominance in at least one of the relevant markets. The effects-based analysis should also take into account any efficiencies or benefits to consumers of the firm's strategy, such as, for example, whether certain product integrations are genuine product improvements. Such a careful assessment of all existing integration strategies by digital players is beyond the scope of this report.

4 CONCLUDING REMARKS

90. GenAI is a transformative technology with the potential for lasting impactful change in all markets, digital and non-digital. It has been suggested that this new technology alone could increase global GDP by 7% in one decade. Given the pace of new developments in the recent past, this prediction may appear, if anything, conservative.

91. Given this potential, it is perhaps unsurprising that competition authorities across the globe show interest in this sector, particularly demonstrating a keen interest in avoiding their self-professed failure to allow for the creation of substantial market power with past digital innovations.

92. Competition authorities have identified risks and uncertainties for future developments arising from the following factors:

- potential lack of access to a scarce set of resources, including data, computing power and hardware, and human talent
- effects of specific partnership agreements between firms at different levels in the value chain
- risk that certain business practices may allow firms to leverage existing market power into new markets (foreclosing behaviour) or exploit any existing market power.

GenAI is a diverse and vibrant sector with no immediate signs of lack of access to inputs.

93. A variety of different firms are already present in the GenAI sector, and the speed of innovation shows no sign of slowing down. These firms are diversified in, among other things, their approach to the market, their size, and their access to capital, ranging from fully integrated players who develop their own foundation models and user-facing applications to highly specialised players who develop specific new applications based on available foundation models.

94. While making predictions in ever-changing digital markets is always a fraught exercise, at this stage, we do not see evidence that access to inputs has barred innovative firms from making a dent in this segment. Most notably, OpenAI, a startup company started in 2015 as a non-profit AI research lab, has quickly become a strong competitor to the GenAI models of more established digital players. However, its position is in no way guaranteed. Besides competition from larger digital players, it also faces competition from new startups such as Mistral, a startup founded less than 12 months ago that has managed to quickly develop a model capable of competing head-to-head with the models already on the market.

95. Uncertainties remain regarding the role of inputs in the long run. In particular, some of the key inputs, such as chips and cloud computing, are produced by very few firms. While a decline on required computational power is likely, models are becoming increasingly complex. The future importance of other inputs, such as data and talent, is uncertain. A promising development is that many components of foundation models (e.g., model weights), including those developed by large players in the digital sector, are publicly available for competing downstream developers to freely fine-tune them and develop new GenAI applications.

Care should be taken in evaluating the impacts of any regulation.

96. Given that GenAI is still a nascent technology, it is also important to ensure that any regulation is fit for purpose and does not lead to a dampening of competitive conduct, including entry and expansion of new players. A survey of smaller firms (which are less capable of bearing the costs of regulation) found that regulation ranks second only to financing as the main barrier for EU startups in the AI segment more generally. A careful further analysis of future and current regulatory interventions should be undertaken to understand the full impact of these interventions on GenAI developers, and whether it is possible to soften the burden on SMEs.

Partnership agreements between large tech players and AI startups serve an important purpose, but need to be carefully assessed.

97. Certain partnership agreements may give large tech firms excessive power over or privileged access to GenAI development startups. We have already seen recently an increased interest from authorities in Europe in the partnership between Microsoft and Open AI. Even if a partnership does not result in a merger situation, agreements with an anticompetitive effect should still be carefully reviewed.

98. The potential anticompetitive impact will depend on a number of factors, including the nature of the partnership, the form of control exercised and the impact on incentives, and the market position of the firms. Additionally, the benefits of the partnership should also be considered.

Risk of anticompetitive conduct in the deployment of GenAI

99. GenAI applications are likely to be introduced through the integration of new features into existing digital services. While this integration can boost competition with existing market leaders, it can also give rise to potential anti-competitive foreclosure via practices such as tying, bundling, or self-preferencing. Specifically, for services acting as gateways for European businesses and users, the Digital Markets Act established a framework. In addition, Article 102 TFEU (or equivalents) provide a framework for assessment of such conducts for other services or conducts not covered by the DMA.

REFERENCES

Air Street Capital, & Benaich, N. (2023). *State of AI Report 2023*. Retrieved from https://www.stateof.ai/

Aleph Alpha. (2023). *Luminous: European AI closes gap to world leaders*. Retrieved from https://aleph-alpha.com/luminous-european-ai-closes-gap-to-world-leaders/

Amazon. (2023). *Amazon and Anthropic Announce Strategic Collaboration to Advance Generative AI*. Retrieved from https://press.aboutamazon.com/2023/9/amazon-and-anthropic-announce-strategic-collaboration-to-advance-generative-ai

appliedAI Institute for Europe. (2023). *Generative AI in the European Startup Landscape 2024*.

Autoridade Concorrência. (2023). *Competition and Generative Artificial Intelligence*.

Bloomberg. (2023). *Generative AI to Become a $1.3 Trillion Market by 2032, Research Finds*. Retrieved from https://www.bloomberg.com/company/press/generative-ai-to-become-a-1-3-trillion-market-by-2032-research-finds/

Bundeskartellamt. (2023). *Bundeskartellamt examined whether partnership between Microsoft and OpenAI was subject to notification obligation under merger control* .

Capgemini. (2024). *Capgemini partners with Mistral AI to spearhead the adoption of new frontier generative AI models*. Retrieved from https://www.capgemini.com/news/press-releases/capgemini-partners-with-mistral-ai-to-spearhead-the-adoption-of-new-frontier-generative-ai-models/

CB Insights. (2023). *Generative AI is churning out unicorns like it's 2021*. Retrieved from https://www.cbinsights.com/research/generative-ai-unicorns-valuations-revenues-headcount/

CMA. (2023a). *AI Foundation Models: Initial Report*.

CMA. (2023b). *CMA seeks views on Microsoft's partnership with OpenAI*. Retrieved from https://www.gov.uk/government/news/cma-seeks-views-on-microsofts-partnership-with-openai

CNBC. (2023). *Bosch, SAP and others back Europe's answer to OpenAI with $500 million of fresh funds*. Retrieved from https://www.cnbc.com/2023/11/06/aleph-alpha-raises-500-million-to-build-a-european-rival-to-openai.html

Council of the European Union. (2024). *Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts*.

Datawrapper. (2024). *Global Generative AI VC investment*. Retrieved from https://www.datawrapper.de/_/5Q5fp/

Dealroom. (2023). *The European Deep Tech Report 2023*.

Engadget. (2024). *Google's Gemini 1.5 Pro is a new, more efficient AI model*. Retrieved from https://www.engadget.com/googles-gemini-15-pro-is-a-new-more-efficient-ai-model-181909354.html?guccounter=1&guce_referrer=aHR0cHM6Ly93d3cuZ29vZ2xlLmNvbS8&guce_referrer_sig=AQAAAKdcUQkWC6cK8Sq3UJMeHMNG35PpVn9NpkcJCIStAC2q9_Srb-1-U7j4BQGfpI2_huJDgOCTeGE4gxS

European Commission , Directorate-General for Competition, Crémer, J., Schweitzer, H., & Montjoye, Y.-A. (2019). *Competition policy for the digital era*.

European Commission. (2024a). *Commission launches calls for contributions on competition in virtual worlds and generative AI*. Retrieved from https://ec.europa.eu/commission/presscorner/detail/en/IP_24_85

European Commission. (2024b). *Commission launches AI innovation package to support Artificial Intelligence startups and SMEs*. Retrieved from https://ec.europa.eu/commission/presscorner/detail/en/ip_24_383

Financial Times. (2023a). *ChatGPT parent OpenAI seeks $86bn valuation*. Retrieved from https://www.ft.com/content/e4ab95c9-5b45-4996-a69e-46075d6428e5

Financial Times. (2023b). *Big Tech outspends venture capital firms in AI investment frenzy*. Retrieved from https://www.ft.com/content/c6b47d24-b435-4f41-b197-2d826cce9532

Financial Times. (2023c). *How Microsoft's multibillion-dollar alliance with OpenAI really works*. Retrieved from https://www.ft.com/content/458b162d-c97a-4464-8afc-72d65afb28ed

Financial Times. (2024a). *OpenAI on track to hit $2bn revenue milestone as growth rockets*. Retrieved from https://www.ft.com/content/81ac0e78-5b9b-43c2-b135-d11c47480119

Financial Times. (2024b). *Microsoft strikes deal with Mistral in push beyond OpenAI*. Retrieved from https://www.ft.com/content/cd6eb51a-3276-450f-87fd-97e8410db9eb

Financial Times. (2024c). *AI is too important to be monopolised*. Retrieved from https://www.ft.com/content/1fda45a2-43e0-4c10-b5fb-b6097e3f5c56
Forbes. (2023). *What Large Models Cost You – There Is No Free AI Lunch*. Retrieved from https://www.forbes.com/sites/craigsmith/2023/09/08/what-large-models-cost-you--there-is-no-free-ai-lunch/?sh=7b62a84f4af7
French Competition Authority . (2024). *Generative artificial intelligence: the Autorité starts inquiries ex officio and launches a public consultation open until Friday, 22 March*. Retrieved from https://www.autoritedelaconcurrence.fr/en/press-release/generative-artificial-intelligence-autorite-starts-inquiries-ex-officio-and-launches
French Competition Authority. (2023). *The General Rapporteur of the Autorité de la concurrence indicates that an unannounced inspection was carried out in the graphics cards sector*. Retrieved from https://www.autoritedelaconcurrence.fr/en/press-release/general-rapporteur-autorite-de-la-concurrence-indicates-unannounced-inspection-was
Goldman Sachs. (2023a). *Generative AI could raise global GDP by 7%*. Retrieved from https://www.goldmansachs.com/intelligence/pages/generative-ai-could-raise-global-gdp-by-7-percent.html
Goldman Sachs. (2023b). *AI investment forecast to approach $200 billion globally by 2025*. Retrieved from https://www.goldmansachs.com/intelligence/pages/ai-investment-forecast-to-approach-200-billion-globally-by-2025.html
Google. (2024a). *Gemma: Introducing new state-of-the-art open models*. Retrieved from https://blog.google/technology/developers/gemma-open-models/
Google. (2024b). *Gemini models*. Retrieved from https://ai.google.dev/models/gemini
Hsieh, C.-Y., Li, C.-L., Yeh, C.-K., Nakhost, H., Fujii, Y., Ratner, A., . . . Pfister, T. (2023). Distilling Step-by-Step! Outperforming Larger Language Models with Less Training Data and Smaller Model Sizes. arXiv preprint arXiv:2305.02301.
Hungarian Competition Authority. (2024). *GVH launches market analysis on the impact of artificial intelligence*.
Katz, M. (2018). *Exclusionary conduct in multi-sided markets*.
LMSYS Org. (2023). *Chatbot Arena: New models & Elo system update*. Retrieved from https://lmsys.org/blog/2023-12-07-leaderboard/
Maslej, N., Fattorini, L., Brynjolfsson, E., Etchemendy, J., Ligett, K., Lyons, T., . . . Perrault, R. (2023). *The AI Index 2023 Annual Report*.
McKinsey & Company. (2023). *Exploring opportunities in the generative AI value chain*.
Medium. (2023a). *Behind the Millions: Estimating the Scale of Large Language Models*. Retrieved from https://towardsdatascience.com/behind-the-millions-estimating-the-scale-of-large-language-models-97bd7287fb6b
Medium. (2023b). *Mixtral 8x7B: Redefining AI with Mixture-of-Experts (MOE) Technology*. Retrieved from https://medium.com/@adamlouly/mixtral-8x7b-redefining-ai-with-mixture-of-experts-moe-technology-20b8f04c18a9
Mistral AI. (2023a). *Mistral 7B*. Retrieved from https://mistral.ai/news/announcing-mistral-7b/https://mistral.ai/news/announcing-mistral-7b/
Mistral AI. (2023b). *Mixtral of experts*. Retrieved from https://mistral.ai/news/mixtral-of-experts/
Mistral AI. (2024a). *Au Large*. Retrieved from https://mistral.ai/news/mistral-large/
Mistral AI. (2024b). *Le Chat*. Retrieved from https://mistral.ai/news/le-chat-mistral/
MLex. (2024). *Comment: Privacy takes stage in US AI litigation against Microsoft, Google, OpenAI; scholars say AI requires data-governance shift*. Retrieved from https://mlexmarketinsight.com/news/comment/comment-privacy-takes-stage-in-us-ai-litigation-against-microsoft-google-openai-scholars-say-ai
Motta, M. (2023). Self-preferencing and foreclosure in digital markets: theories of harm for abuse cases. *International Journal of Industrial Organization*(102974).
Motta, M., & Fumagalli, C. (2024). *Economic principles for the enforcement of abuse of dominance provisions*.
NTIA. (2024). *Dual Use Foundation Artificial Intelligence Models with Widely Available Model Weights*.
OpenAI. (2023a). *GPT-4*. Retrieved from https://openai.com/research/gpt-4
OpenAI. (2023b). *Introducing ChatGPT Enterprise*. Retrieved from https://openai.com/blog/introducing-chatgpt-enterprise
OpenAI. (2023c). *Introducing ChatGPT Team*. Retrieved from https://openai.com/blog/introducing-chatgpt-team

OpenAI. (2024a). *Introducing the GPT Store*. Retrieved from https://openai.com/blog/introducing-the-gpt-store
Reuters. (2023a). *ChatGPT sets record for fastest-growing user base*. Retrieved from https://www.reuters.com/technology/chatgpt-sets-record-fastest-growing-user-base-analyst-note-2023-02-01/
Reuters. (2023b). *Google agrees to invest up to $2 billion in OpenAI rival Anthropic*. Retrieved from https://www.reuters.com/technology/google-agrees-invest-up-2-bln-openai-rival-anthropic-wsj-2023-10-27/
Schrepel, T., & Pentland, A. (2023). *Competition Between AI Foundation Models: Dynamics and Policy Recommendations*.
Seger, E., Dreksler, N., Moulange, R., Dardaman, E., Schuett, J., Wei, K., . . . Anderljung, M. (2023). *Open-sourcing highly capable foundation models: An evaluation of risks, benefits, and alternative methods for pursuing open-source objectives*. arXiv preprint arXiv:2311.09227.
Shen, Y., Wallis, P., Allen-Zhu, Z., Li, Y., Wang, S., Wang, L., . . . Hu, E. (2021). LoRA: Low-Rank Adaptation of Large Language Models. *arXiv preprint arXiv:2106.09685*.
Sifted. (2023a). *'The economics of trading equity for compute are not great' — Mistral releases its first model*. Retrieved from https://sifted.eu/articles/mistral-releases-first-ai-model
Sifted. (2023b). *'EU's AI act could kill our company,' says Mistral's Cédric O*. Retrieved from https://sifted.eu/articles/eu-ai-act-kill-mistral-cedric-o
Solaiman, I. (2023). *The gradient of generative AI release: Methods and considerations*.
Stanford University. (2024a). *Ecosystem Graphs for Foundation Models*. Retrieved from https://crfm.stanford.edu/ecosystem-graphs/index.html?mode=table
Stanford University. (2024b). *Rethinking Privacy in the AI Era*.
Tech Crunch. (2023). *OpenAI inks deal with Axel Springer on licensing news for model training*. Retrieved from https://techcrunch.com/2023/12/13/openai-inks-deal-with-axel-springer-on-licensing-news-for-model-training/
Tech Crunch. (2024). *Microsoft made a $16M investment in Mistral AI*. Retrieved from https://techcrunch.com/2024/02/27/microsoft-made-a-16-million-investment-in-mistral-ai/
The Economist. (2024a). *Wannabe AI superpowers seek a homegrown edge*. Retrieved from https://www.economist.com/business/2024/01/01/welcome-to-the-era-of-ai-nationalism
The Economist. (2024b). *Meet the French startup hoping to take on OpenAI*. Retrieved from https://www.economist.com/business/2024/02/26/meet-the-french-startup-hoping-to-take-on-openai
The Information. (2024a). *Microsoft forms team to make cheaper Generative AI*. Retrieved from https://www.theinformation.com/articles/microsoft-forms-team-to-make-cheaper-generative-ai
The Information. (2024b). *Anthropic's Antitrust Advantage*. Retrieved from https://www.theinformation.com/articles/anthropics-antitrust-advantage
The New York Times. (2023). *Mistral, French A.I. Start-Up, Is Valued at $2 Billion in Funding Round*. Retrieved from https://www.nytimes.com/2023/12/10/technology/mistral-ai-funding.html#:~:text=The%20deal%20values%20the%2022,familiar%20with%20the%20deal%20said.
The Verge. (2023). *Microsoft is finally making custom chips — and they're all about AI*. Retrieved from https://www.theverge.com/2023/11/15/23960345/microsoft-cpu-gpu-ai-chips-azure-maia-cobalt-specifications-cloud-infrastructure
TheNorthAI. (2023). *Bosch Unveils 'BoschGPT' for Employee Communication by Year-End*. Retrieved from https://thenorth.ai/2023/08/20/bosch-unveils-boschgpt-for-employee-communication-by-year-end/
Wired. (2023a). *Stack Overflow Will Charge AI Giants for Training Data*. Retrieved from https://www.wired.com/story/stack-overflow-will-charge-ai-giants-for-training-data/
Wired. (2023b). *OpenAI's CEO Says the Age of Giant AI Models Is Already Over*. Retrieved from https://www.wired.com/story/openai-ceo-sam-altman-the-age-of-giant-ai-models-is-already-over/