Organisation for Economic Co-operation and Development

DAF/COMP/WD(2025)74

Unclassified

English - Or. French

28 November 2025

DIRECTORATE FOR FINANCIAL AND ENTERPRISE AFFAIRS
COMPETITION COMMITTEE

Competition in Artificial Intelligence Infrastructure – Note by France

5 December 2025

This document reproduces a written contribution from France submitted for Item 7 of the 147th OECD Competition Committee meeting on 4-5 December 2025.

Antonio CAPOBIANCO
Antonio.Capobianco@oecd.org, +(33-1) 45 24 98 08

JT03577804

France

1. On 28 June 2024, the *Autorité de la concurrence* (the *Autorité*) published Opinion 24-A-05 on the competitive functioning of the generative artificial intelligence sector[1] (hereafter generative AI). The opinion aims to provide stakeholders with a competitive analysis of the fast-growing generative AI sector.

2. During the investigation carried out for its opinion, the *Autorité* launched a public consultation during which more than 40 stakeholders and a dozen stakeholder associations expressed their views. At the same time, the Autorité held interviews with French and international private players (suppliers, customers, associations and others) and institutional players (government departments, Industry Regulators, foreign competition authorities, etc.). In particular, the *Autorité* exchanged views with the authorities that conducted in-depth work on the competition issues raised by the generative AI sector: the Portuguese competition authority (*Autoridade da concorrência*), the European Commission, the US *Federal Trade Commission* (FTC) and the UK *Competition and Markets Authority* (CMA).

3. Generative AI has made a remarkable and rapid entry into the public arena since the launch of version 3.5 of the ChatGPT chatbot by the US company OpenAI in November 2022. This technology represents a major technological revolution, distinguished by its ability to generate new content, an unprecedented rate of adoption - 100 million users in two months for ChatGPT2 - and massive investment. This sharp uptake has catalysed interest among the public and companies, making generative AI a central subject of public and economic debate.

4. The *Autorité's* opinion had a particular focus on the strategies implemented by major digital companies aimed at consolidating their market power upstream in the generative AI value chain, i.e. the design, training and fine-tuning of large language models (LLMs) or at leveraging this market power in order to expand in this booming sector. It also took an interest in analysing minority shareholdings in the sector.

5. This contribution will first present the different players in the generative AI value chain, before setting out the lessons learned from the *Autorité's* opinion specific to computing power, a key input for the development of foundation models.

6. In conjunction with the *Autorité*, in September 2025, as part of the review process of the Digital Markets Act (DMA), the French authorities presented several proposals toward the evolution of the DMA. In line with the proposals made by the French authorities and continuing this work, while France and Germany, at the European Summit on digital sovereignty, on 18 November 2025, called for greater contestability and fairness in the cloud and AI markets, the Commission has just opened two investigations under the DMA into the cloud market, which represents a critical competitive infrastructure for AI.

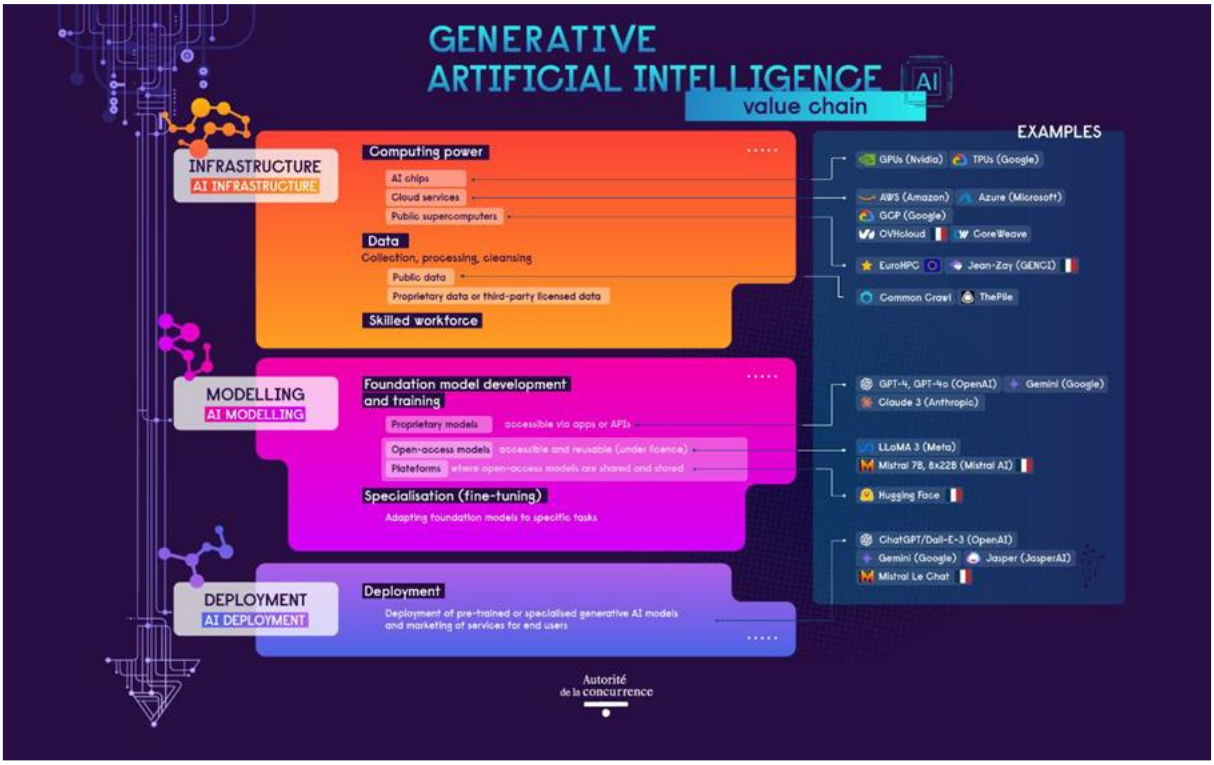
[1] https://www.autoritedelaconcurrence.fr/en/press-release/generative-artificial-intelligence-autorite-issues-its-opinion-competitive

[2] https://www.itforbusiness.fr/chatgpt-depasse-deja-les-100-millions-dutilisateurs-mensuels-59523

1. The different players in the generative AI value chain

7. In its opinion, the *Autorité* outlined the competitive landscape of the generative AI sector. Although this sector is developing very rapidly, this presentation is still valid on the date of this contribution, and a summary is presented below.

Figure 1. The main operators in the generative AI value chain



Source: Autorité de la concurrence

1.1. The major digital companies in the sector

8. The major digital companies seem to have adopted different strategies for generative AI. Some (Alphabet and Microsoft) are present across the entire value chain (vertical and conglomerate integration), while others (Amazon, Apple, Meta and Nvidia) are present only at certain specific layers.

1.1.1. Alphabet (Google)

9. Alphabet provides cloud computing services via Google Cloud Platform (GCP). It has developed its own tensor processing units (TPU), optimised for the training and inference of large AI models and offered to GCP customers. Alphabet owns the largest search index (Google Search), as well as the world's largest video database with YouTube. The group has developed the Gemini (TPU-based training) and Gemma foundation models. It offers the Gemini chatbot (formerly Bard).

1.1.2. Amazon

10. Amazon provides cloud services via its subsidiary Amazon Web Services (AWS). AWS provides specialised AI chips called "Trainium" (for training) and "Inferentia" (for

inference). Amazon is also positioning itself as a model developer, via its Titan model range, which is accessible via its Amazon Bedrock platform. Amazon integrates generative AI systems into its products, such as the Alexa voice assistant.

1.1.3. Apple

11. Apple specialises in the design, manufacture and marketing of electronic products (iPhone, iPad and Mac) and software. The company has presented its first models developed in-house, including the large proprietary MM1 model and a range of small open-source models, OpenELM. Apple has announced the launch of generative AI-based features called “Apple Intelligence” in the latest versions of its products.

1.1.4. Meta

12. Meta does not provide cloud services, but has large on-site IT infrastructures for operating its platforms (Facebook, Instagram, WhatsApp). It develops and designs its own AI accelerators. Its research laboratories have developed the Llama range of language models (Llama 2, Llama 3), which are generally published as open weights (subject to certain licensing conditions). Meta has launched Meta AI, a chatbot based on the Llama models.

1.1.5. Microsoft

13. Microsoft is one of the world's leading cloud service providers. It has announced the release of AI accelerators called Maia. It is a partner of OpenAI and integrates their models into its products, while developing its own small language models (SLMs) such as Phi (including Phi-3, published as open weights). Microsoft integrates generative AI into its historical services via the "Copilot" function (for Microsoft Bing and Microsoft 365).

1.1.6. Nvidia

14. Nvidia specialises in the development and design of graphics processing units (GPUs), which are essential for model training. It was the world’s most valuable publicly traded company at the date of this opinion. Nvidia is known for its proprietary CUDA software, which is the most widely used environment for programming on its GPUs. It has extended its presence upstream by forging partnerships with a number of cloud service providers (AWS, GCP, Microsoft Azure).

1.2. Developers of generative AI models

15. These are operators, often research laboratories or innovative companies native to AI, who design and train the foundation models. They include:

1.2.1. OpenAI

16. First to offer its generative AI model to the general public via the chatbot ChatGPT. It is developing GPT models, as well as image generation (DALL-E) and video generation (Sora) models.

1.2.2. Mistral AI

17. *French* start-up that develops models published as *open weights* or proprietary formats and power its Le Chat chatbot.

1.2.3. Anthropic

18. Research laboratory that developed a range of large language models marketed under the name Claude.

1.2.4. HuggingFace

19. French-US company that provides a hosting and collaboration platform for the vast majority of open-source and open-weights models.

20. Other operators, such as Cohere, LightOn, Stability AI and xAI, are also developing their own foundation models.

1.3. The operators upstream in the value chain

21. Upstream in the value chain is crucial for model training and includes several essential inputs.

1.3.1. IT component suppliers

22. These operators develop graphics processing units and AI accelerators, which are essential for training. In addition to Nvidia, historical competitors include Advanced Micro Devices (AMD) and Intel. New entrants such as Cerebras, Graphcore, SambaNova and Groq are moving into this sector.

1.3.2. Cloud service providers

23. They provide the storage, data processing, and computing capabilities needed by model developers. Hyperscalers (AWS, GCP, Microsoft Azure) are the main providers.[3] Other companies, such as OVHcloud, IBM and Scaleway, also provide services. Specialised operators such as the US company CoreWeave, a preferred partner of Nvidia, focus on providing high-performance computing services for AI.

1.3.3. Public supercomputers:

24. These "very large computers" have traditionally been used for high-performance computing and have diversified to accommodate AI research projects. They provide computing resources to researchers free of charge. A French example is the Jean Zay supercomputer. The joint European initiative EuroHPC is also working to develop these supercomputers across Europe and plans to install a new supercomputer in France.

1.4. The main operators downstream in the value chain

25. Downstream operators offer generative AI-based products and services to end users.

1.4.1. Major technology companies integrating generative AI tools:

26. These operators are integrating AI tools into their existing products and services. Examples include Adobe (with Firefly in Photoshop), Samsung (in its Galaxy S24 smartphone range) and Zoom (with Zoom AI Companion).

[3] A term for very large companies that have built global hosting capabilities and developed dedicated applications used by millions of users.

1.4.2. Providers of products and services for users, companies and the public sector:

27. Most model developers offer an internet interface, most often in the form of chatbots such as ChatGPT (OpenAI), Gemini (Google) or Le Chat (Mistral AI).

28. Several hundred companies also offer specialised applications based on generative AI for companies in different sectors (sales, marketing, HR, etc.).

29. Lastly, digital service companies such as Accenture, Atos and Capgemini, and start-ups, such as Dust and AleIA, offer facilitator services on behalf of their customers.

2. A sector characterised by high barriers to entry: the crucial aspect of computing power

30. The generative AI sector is characterised by high barriers to entry. Access to computing power is a major input in the development of foundation models.

2.1. The need for specialised, expensive equipment

31. To develop a foundation model, access to sufficient computing power is essential for performing a large number of operations in parallel, and with the high precision needed to determine the several billion parameters that make up the model.

32. The graphics processing units (GPUs) developed in particular by Nvidia, or the AI accelerators (such as the TPUs [tensor processing units] of Google and the chips developed by AWS and Microsoft) are essential for the training, fine-tuning and inference of generative AI models. They offer greater computing power than central processing units (CPUs), being capable of performing several thousand mathematical operations in parallel.

33. However, these chips are very expensive. For example, Nvidia's H100 GPUs cost between €30,000 and €40,000 each. Training a foundation model requires several thousand of these processors. Training the GPT-4 model alone cost over $78 million (around €72 million).

34. Furthermore, on the date of the opinion, the sector was experiencing difficulties in sourcing GPUs and other AI accelerators due to an explosion in demand and a limited supply of semiconductors, a situation still present in 2025.

35. In addition, as research within GPUs is extremely dynamic, this hardware can quickly become obsolete, which increases the constraints and costs involved in maintaining an on-site infrastructure.

36. Lastly, in addition to the hardware, model creation also requires a software layer. The proprietary CUDA software environment developed by Nvidia is the most widely used by operators in the sector, creating a fundamental dependence and interplay between the hardware and the software that reinforces mergers and inertia.

2.2. The cloud: the key to computing power

37. Faced with massive costs and constraints linked to specialised equipment, the use of cloud services has become the most common way to access the computing power needed for model training.

38. This type of hardware can be very expensive. In addition to the initial investment, the costs of use (electricity, cooling) and implementation (premises to be fitted out) are

significant. As a result, only a few companies have sufficient on-site infrastructure to develop generative AI models.

39. The cloud offers the advantage of avoiding initial investment and maintenance costs. It gives developers access to AI-specific infrastructure and platform services while allowing for pay-as-you-go pricing and flexibility based on business needs.

40. Hyperscalers (AWS, GCP, Microsoft Azure) play a dual and strategic role as they are an essential gateway for model training and fine-tuning. They are also a privileged platform for the deployment of foundation models (Model as a Service, MaaS), reinforcing their role as intermediaries for downstream distribution.

2.3. Alternatives for limiting barriers to entry

41. The opinion lists several solutions that could reduce barriers to entry related to computing power.

42. **Public supercomputers**, traditionally focused on high-performance computing, are accessible free of charge to certain users, particularly in the research community, in exchange for a contribution to open science (e.g. publication of work). The French Jean Zay supercomputer is an example used for model training. However, these supercomputers are not a solution for inference (content production) as they are only accessible for a limited time.

43. Innovations in model architecture, such as **Mixture of Experts (MoE)**[4] or **Low-Rank Adaptation (LoRA)**[5], also make training more efficient and less costly by reducing computing power requirements**.**

44. Furthermore, since the opinion was published, Chinese operators have made a notable appearance, in particular with the "DeepSeek moment", corresponding to the release in January 2025 of the open-source model DeepSeek-R1[6]. Designed at low cost while achieving a performance level comparable to the most advanced models, it has confirmed the relevance of the open-source approach over proprietary solutions. It is also notable that most developers publish both proprietary and open-source models. OpenAI, which had not released a free model since GPT-2, recently made two open-source models available to the public.[7]

45. Lastly, **the existence of open models** (open weights), generally through the publication of their weights, can reduce barriers to entry for fine-tuning and inference. However, it does not reduce the barrier to entry for new operators wanting to train foundation models. Knowledge of these weights is of only marginal use for training foundation models, the purpose of which is to create new models and their own weights.

[4]MoE/Mixture of Experts: AI model architecture divided into several subsets of neural networks called experts, specialised in a specific task, and a router that determines which expert should be used to answer a query.

[5]LoRA/Low Rank Adaptation: fine-tuning technique introduced by a team of Microsoft researchers in 2021, requiring less computing power.

[6] "Intelligence artificielle : le 'moment DeepSeek' de l'économie chinoise" (*Artificial intelligence: the "DeepSeek moment" of the Chinese economy*), Le Monde, 22/2/2025.

[7] "OpenAI renoue avec l'ouverture avec deux modèles open-weight : GPT-OSS-120B et GPT-OSS-20B" (OpenAI returns to openness with two open-weights models: GPT-OSS-120B and GPT-OSS-20B), ActuaIA, 05/09/2025

3. Competition risks linked to computing power

46. In its opinion, the *Autorité* identified several major competition risks linked to access to and control of computing power. These risks can be found at both the component and cloud service levels.

3.1. Competition risks linked to components

47. The components sector (graphics processing units and AI accelerators) is characterised by high concentration, which may raise competition concerns.

48. A single operator, Nvidia, appears to have a prominent position in this sector. This situation could encourage the implementation of potentially anticompetitive practices, such as price fixing, supply restrictions, unfair contractual conditions and discriminatory behaviour.

49. Stakeholders have also expressed concerns regarding the sector's dependence on Nvidia's CUDA chip programming software.

50. Furthermore, Nvidia's recent investments in AI-focused cloud service providers such as CoreWeave are raising concerns regarding GPU access.

51. This graphics card sector, which was the target of a dawn raid in September 2023, is being closely monitored by the *Autorité's* Investigation Services.[8]

3.2. Competition risks linked to the cloud

52. Hyperscalers play an important role in providing the computing resources required by model developers, but their position creates a potential risk of abuse aimed at expanding their market power and reducing competition.

53. The opinion shows that particularly high levels of cloud credit are being offered, especially to innovative generative AI start-ups. These practices could have the effect of locking the companies concerned into hyperscaler ecosystems, particularly against a backdrop of technical and price barriers to migration.

54. Technical lock-in practices have also been identified. Hyperscalers may offer proprietary solutions (e.g. automated machine learning services) that do not give users access to the final model itself. As a result, if the user wanted to change cloud service provider, they would have to recreate their AI model from scratch, effectively locking them in.

55. Exclusivity agreements between cloud service providers and foundation model developers could ultimately lead to competition risks.

3.3. Competition risks associated with the vertical integration of major digital operators

56. Major digital companies (notably Alphabet and Microsoft) enjoy preferential and easier access to computing power as partners and competitors of AI chip suppliers. They are able to buy GPUs in large quantities and negotiate preferential agreements with suppliers such as Nvidia, which gives them access to resources even in times of scarcity.

[8] *Autorité* press release, The General Rapporteur of the *Autorité de la concurrence* indicates that an unannounced inspection was carried out in the graphics cards sector, 27 September 2023.

57. Most major players are also developing their own AI accelerators in-house (Google's TPUs, AWS' Trainium, Microsoft's Maia), which means they are not dependent on the product development cycles of external suppliers. Optimising their own chips for their own data centres is a significant competitive advantage.

58. Upstream, model developers could be denied or given limited access to the chips or data needed to train competing foundation models, which would lead to delays or the implementation of less ambitious models, undermining effective competition.

59. Lastly, there is a risk that major digital companies' preferential access to key inputs, combined with the advantages they enjoy due to their vertical and conglomerate integration, could create the conditions for significant concentration to their benefit.

4. The *Autorité's* recommendations and outlook

60. Access to computing power is essential for the development of research and the emergence of new companies in the generative AI sector. The French authorities highlight in particular the following two recommendations by *Autorité* on this topic:

- continuing to invest in the **development of supercomputers at European level**, to give as many parties as possible access to computing power (Recommendation 5);
- proposing an **open, non-discriminatory framework** that would enable companies to use public supercomputer resources for a fee, while maintaining priority for research, particularly academic research (Recommendation 6).

61. Since the publication of its opinion, the *Autorité* has also noted that announced investments in AI, and more specifically in data centres, have intensified: from around €20 billion in 2023 to several hundred billion today. One example is OpenAI's Stargate project, an ambitious infrastructure programme with an estimated cost of €400 billion[9].

62. At the AI summit held in Paris in 2025, President Emmanuel Macron announced that €109 billion in private investment had been committed to the AI sector in France[10]. At the same time, European Commission President Ursula von der Leyen launched the InvestAI initiative, which aims to mobilise €200 billion in investment in AI[11].

63. For its part, the *Autorité* will remain particularly vigilant with regard to competition issues in the sector. It has already announced that this opinion will be followed by a study on the competitive challenges of energy access for operators in this sector[12].

[9] "Stargate, le projet à 500 milliards de dollars de Donald Trump pour doper l'intelligence artificielle" (Stargate, Donald Trump's $500 billion project to boost artificial intelligence), *Le Monde*, 22/01/2025.

[10] "Intelligence artificielle : Emmanuel Macron annonce des investissements en France de "109 milliards d'euros dans les prochaines années"" (Artificial intelligence: Emmanuel Macron announces investments in France of '€109 billion over the next few years'), *Le Monde*, 09/02/2025.

[11] European Union Press Release, *EU launches InvestAI initiative to mobilise €200 billion of investment in artificial intelligence*, 11/02/2025.

[12] Autorité de la concurrence Roadmap 2025/2026

5. With a view to strengthening the existing framework for the AI market, in November 2025 the Commission launched two investigations into the cloud market

64. At the European Summit on digital sovereignty, France and Germany called for greater contestability and fairness in the cloud and AI markets. In this context, and in line with the recommendations made by the French authorities in their note of September 2025, produced in response to the Commission's consultation with DMA Committee members, the Commission announced on 18 November 2025 the opening of two DMA investigations aiming to:

- determine, within 12 months, whether two core platform services in the "cloud computing services" category (Amazon AWS and Microsoft Azure) should be designated as "important gateways" under the DMA;
- examine, within 18 months, the need to adopt amendments to better address anti-competitive practices in the cloud sector (barriers to interoperability, limited access to data, tying and bundling practices, etc.).

65. These prospects for subjecting two major players to regulation and extending the obligations of the DMA represent tangible levers for addressing the competitive risks associated with the cloud in the generative AI market.