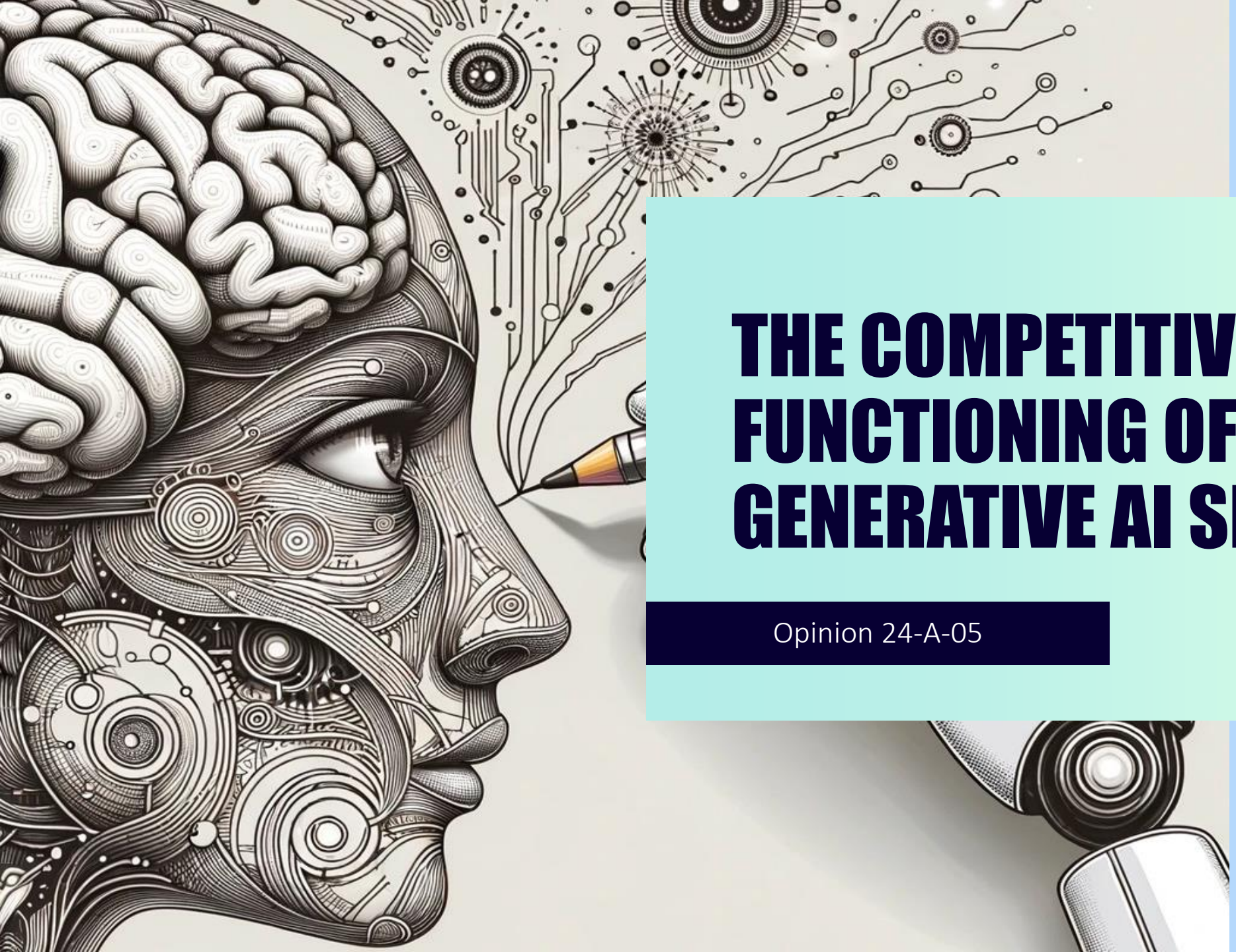


# THE COMPETITIVE FUNCTIONING OF THE GENERATIVE AI SECTOR

Opinion 24-A-05

Autorité de la concurrence

# DEFINITION

- Artificial intelligence refers to any tool used by a machine "*to display human-like capabilities such as reasoning, learning, planning and creativity*" (EU Parliament definition).

- Generative AI refers to AI models capable of generating new content such as text, image, sound or video.

# A GROWING PRIORITY FOR PUBLIC AUTHORITIES

Several initiatives:

- France: national AI strategy, AI Commission.
- Europe: AI Act, DMA, Data Act.
- Global: G7 priority, Bletchley Declaration.

**The benefits of generative AI will only materialise if all households and companies have access to a variety of different models adapted to their needs.**

# TIMELINE

- **8 February 2024:** launch of *ex officio* inquiries and a public consultation:
  - Responses from around 40 stakeholders and 10 associations to the questionnaire.
  - Meetings with private and global operators in the sector and various institutional parties (ministerial departments, sector-specific regulators, European Commission, international competition authorities).

- **29 May 2024**: hearing of the *Autorité*.

# OBJECTIVES AND SCOPE

## SCOPE:

The opinion examines the strategies implemented by major digital companies aimed at consolidating their market power at the upstream level of the value chain or at leveraging that market power in order to expand in the booming generative AI sector.

The *Autorité* focuses in particular on practices implemented by operators already present in cloud infrastructure and at issues relating to access to cloud infrastructure, computing power, data and skilled workforces.

The *Autorité* also examines investments and partnerships by major digital companies, in particular in innovative companies specialised in generative AI.

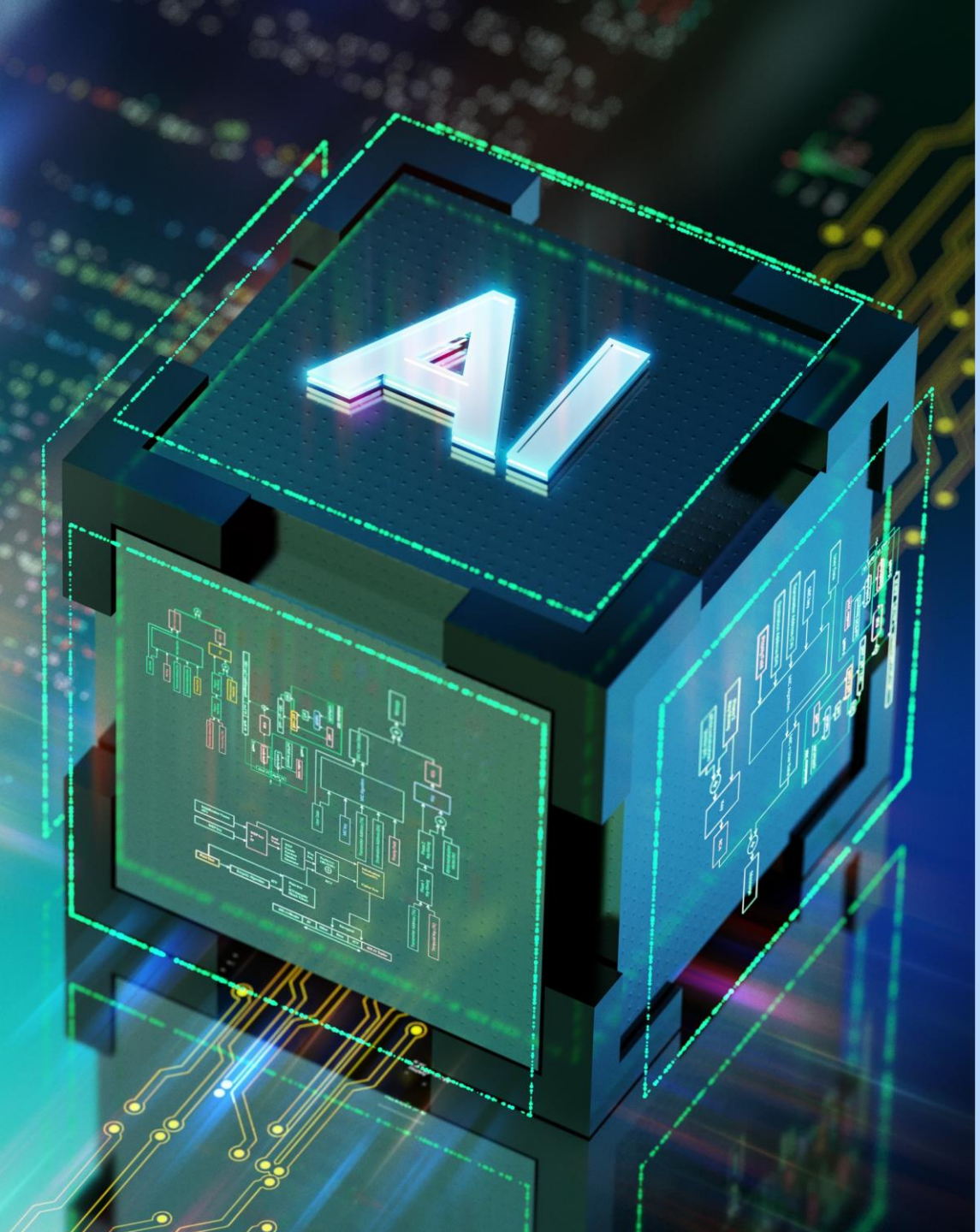
## OBJECTIVES:

- Provide stakeholders with an analysis of the competitive functioning of the generative AI market.
- Make recommendations to foster competition in the sector.

# ADDED VALUE OF THE OPINION

- Analysis of key issues for the future:
  - Agreements with content creators on remuneration.
  - Competitive assessment of minority investments by major digital companies in innovative start-ups.
  - Removal of some of the barriers to entry through open-source technology.
  - Practices that may be implemented in the labour market.

- Vigilance with regard to:
  - Practices that unreasonably restrict access to essential inputs.
  - Partnerships by already dominant digital companies, with or without exclusivity clauses.
  - Practices of tying likely to consolidate, in a lasting way, the generative AI sector around already dominant digital companies.




# PRESENTATION OUTLINE

1. **The generative AI sector**
   - The value chain
   - The participants in the value chain

2. **The competitive functioning of the sector**
   - High barriers to entry
   - The advantages of certain operators due to their presence on other markets
   - The competition risks

3. **The recommendations made by the *Autorité***

PART 1

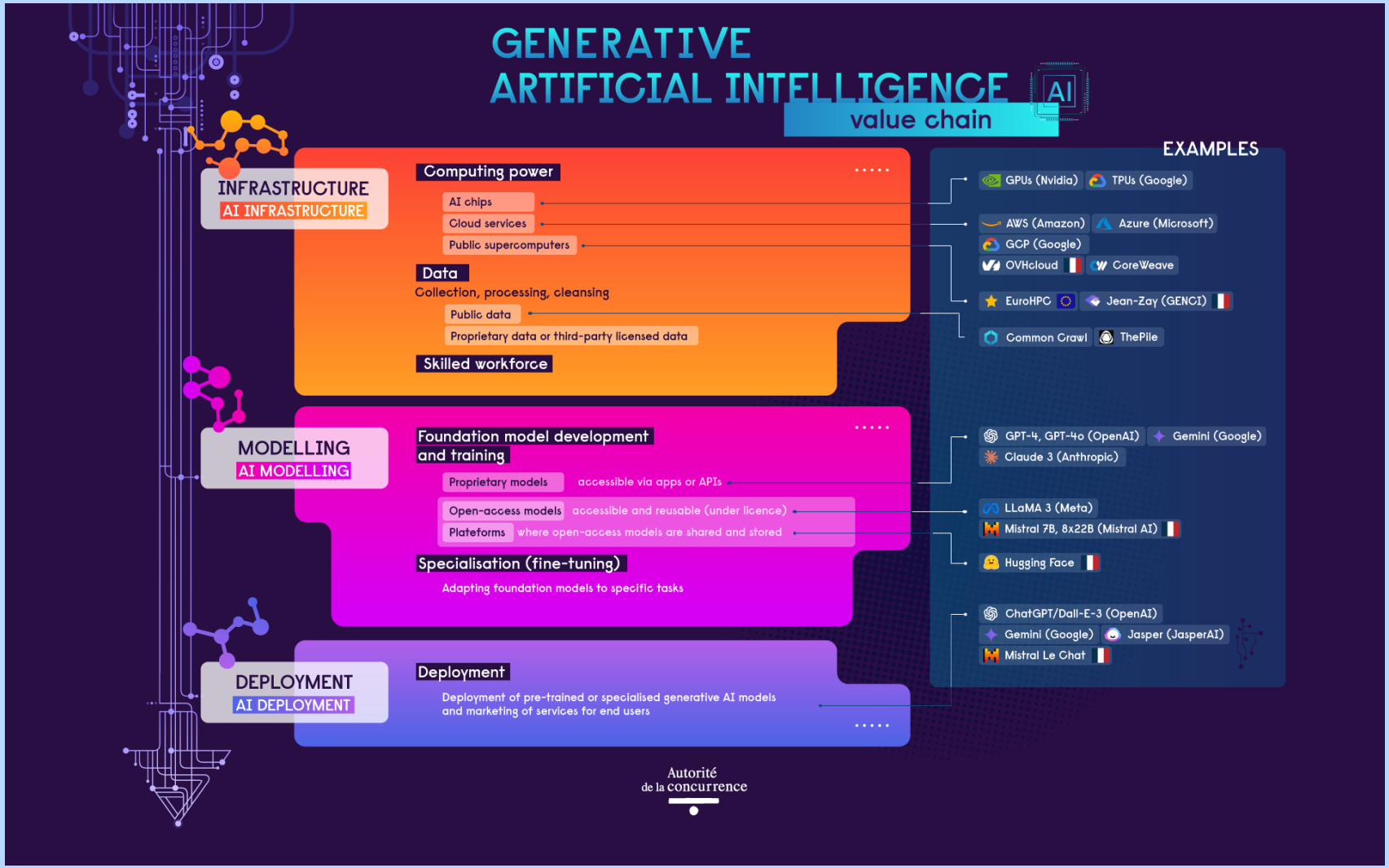
# The generative AI sector

# THE VALUE CHAIN

There are two key phases in generative AI modelling:

- **Training:** the initial learning process of a model, during which its parameters, known as “weights”, are determined. Training requires both significant computing power and a large volume of – generally public – data.
    - The training phase may be followed by **fine-tuning**, during which the model is adapted to a specific task. Fine-tuning is generally based on a smaller, proprietary dataset and may involve human expertise.
- **Inference:** the use of the trained model to generate content. The model can be made accessible to users via specific applications (Mistral AI’s Le Chat) or APIs for developers. The computing power required depends on the number of users.
    - New data that was not used for training may be added during the inference phase, in order to ground the model in recent data, such as news articles.

# THE GENERATIVE AI VALUE CHAIN



# THE PARTICIPANTS IN THE VALUE CHAIN

- **Digital giants:** Alphabet and Microsoft are present across the entire value chain (vertical and conglomerate integration), while Amazon, Apple, Meta and Nvidia are present only at certain specific layers.

- **Model developers** such as Anthropic, HuggingFace, Mistral AI and OpenAI. They have often formed partnerships with one or more digital giants.

At the upstream level, several types of operators are involved:

- **IT component suppliers**, which develop graphics processing units (GPUs) and AI accelerators. Nvidia is the sector's leading operator.
- **cloud service providers**, including "hyperscalers" such as Amazon Web Services (AWS), Microsoft Azure and Google Cloud Platform (GCP), cloud providers such as OVHCloud, as well as specialist AI providers such as CoreWeave.

At the downstream level, many operators are marketing new services based on generative AI to the general public (e.g. ChatGPT), companies and public authorities and/or integrating generative AI into their existing services (e.g. Zoom).

PART 2

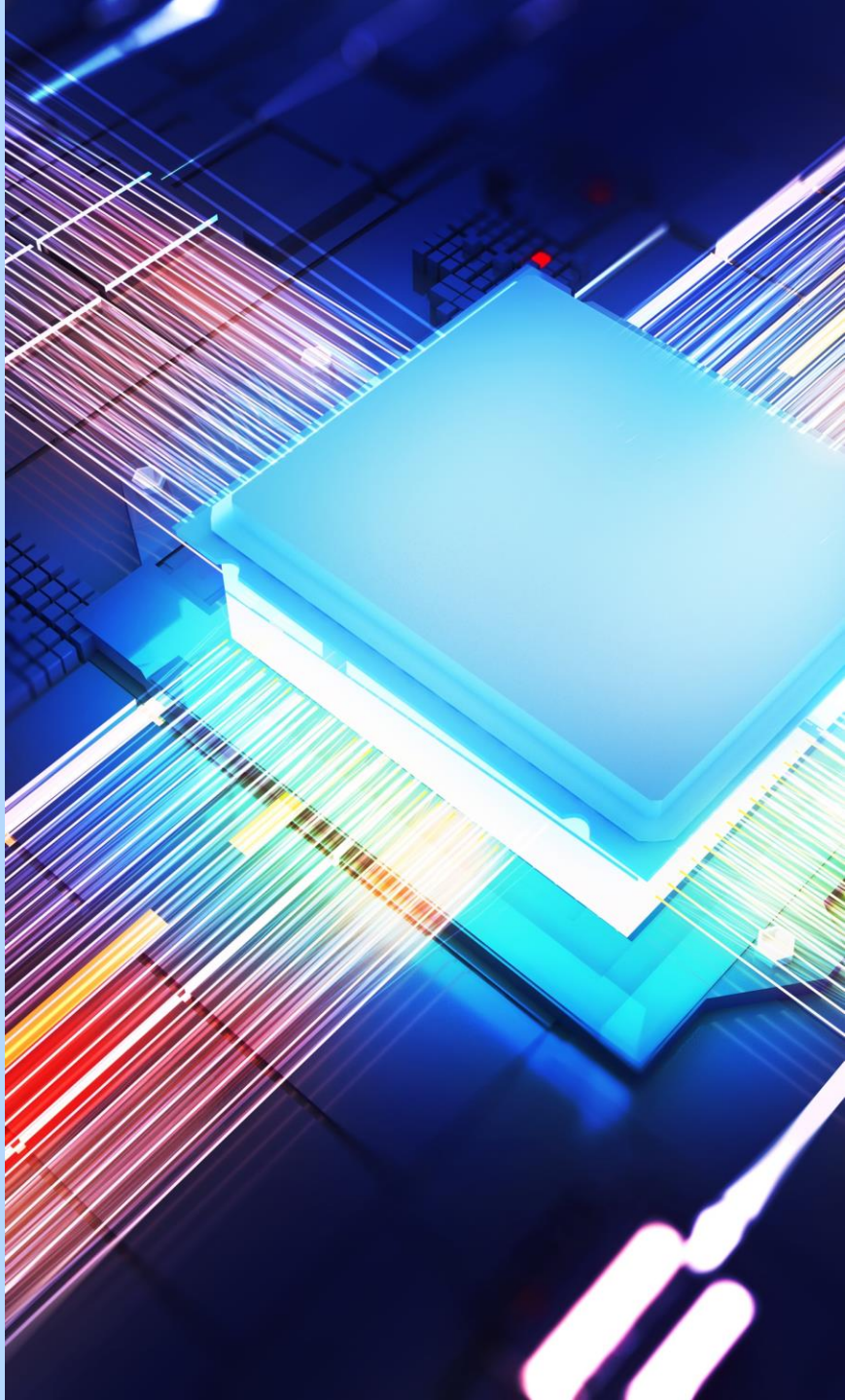
# The competitive functioning of the sector

# HIGH BARRIERS TO ENTRY

- Access to sufficient, high-precision computing power for performing a large number of operations in parallel is essential;
- The cloud is the preferred solution for training and fine-tuning models and also facilitates downstream deployment.
- Model training requires large volumes of data.
- The skills needed are rare and highly sought-after.
- There was a six-fold increase in financing needs between 2022 and 2023.

# FACTORS LIKELY TO LIMIT THE BARRIERS TO ENTRY

- Public supercomputers (e.g. Jean Zay in France).
- Technological innovations that can reduce data and computing power needs (e.g. synthetic data).
- Open-source (or open-weights) models.



# THE ADVANTAGES OF CERTAIN OPERATORS DUE TO THEIR PRESENCE ON OTHER MARKETS

### *Preferential access to inputs*

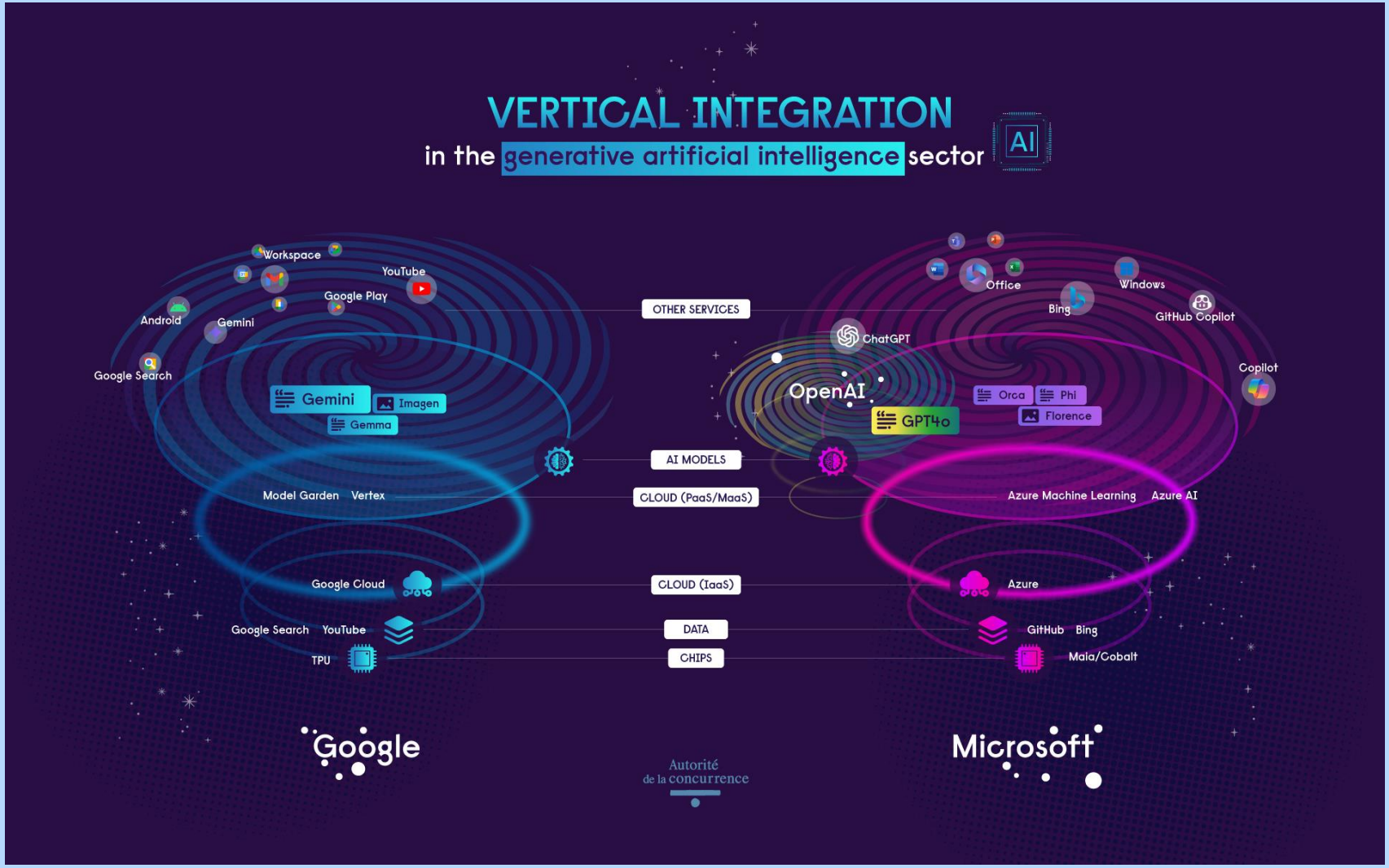
Major digital companies enjoy preferential access to the inputs needed to train and develop foundation models:

- **Easier access to computing power**.
- **Preferential access to large volumes of data** (e.g. Alphabet with YouTube). They can also use their financial power to enter into agreements with content providers.
- **Many highly-skilled employees are enticed by the attractive salaries and employment conditions** offered by major digital companies.

### *The benefits of vertical and conglomerate integration*

- **Economies of scale and scope and network effects.**
- Major digital companies are starting to **integrate generative AI tools into their product and service ecosystems** (Microsoft Copilot).
- Major digital companies' marketplaces **(MaaS)** provide access to proprietary and third-party generative AI models designed to run in their cloud ecosystems.

# THE POSITION OF CERTAIN OPERATORS IN RELATED MARKETS



# THE COMPETITION RISKS (1)

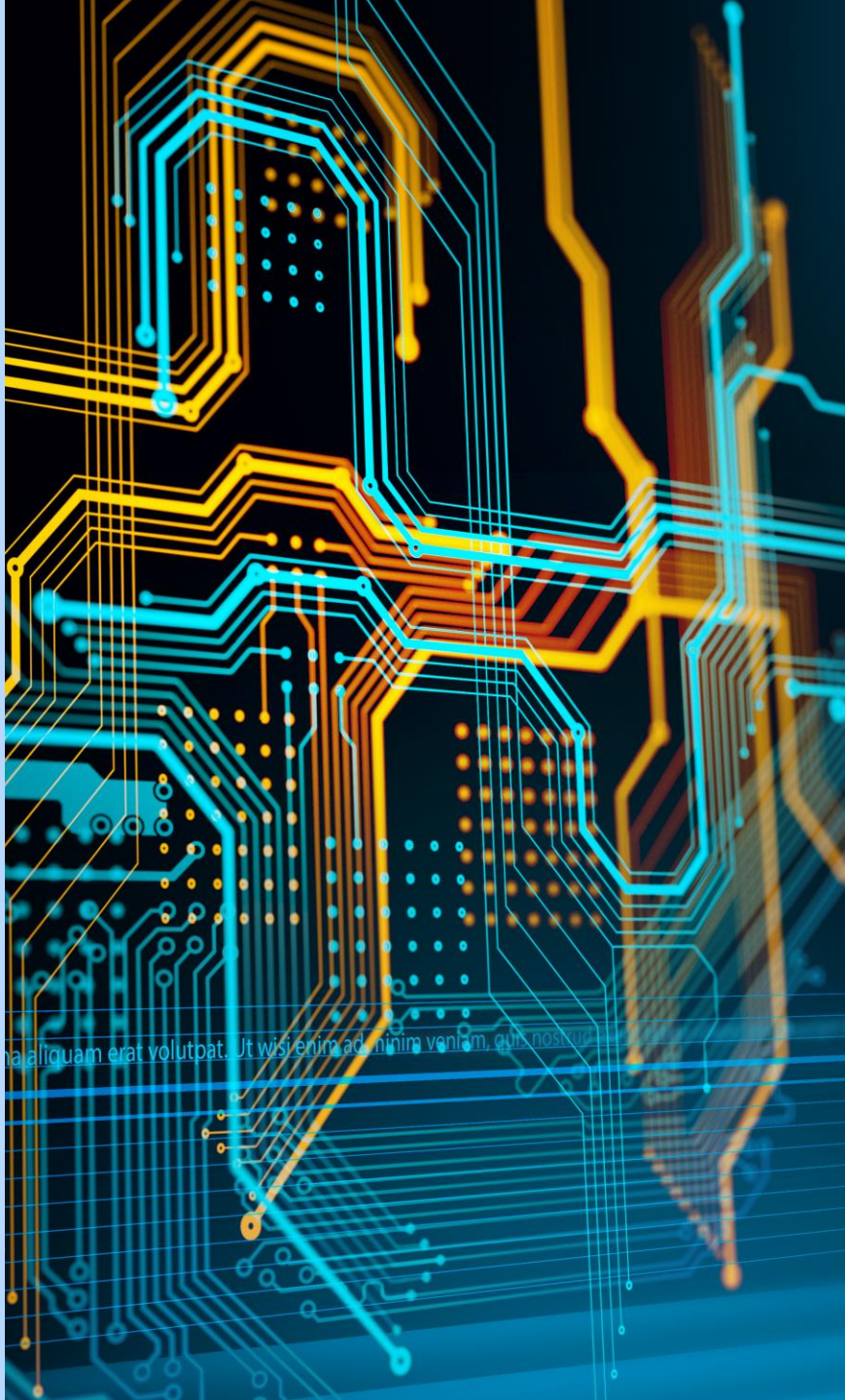
It is difficult, at this stage, to define the **relevant markets**. In some cases, the analysis will have to be conducted in terms of **ecosystems**, rather than market by market.

The *Autorité* identified several risks of abuse:

- **Chips:** price fixing, supply restrictions, unfair contractual conditions and discriminatory behaviour. Concerns regarding the CUDA software.
- **Lock-in by major cloud service providers:** high levels of cloud credits offered specifically to innovative start-ups and technical lock-in practices.

The *Autorité* identified **several competition concerns regarding data access**, in particular:

- Agreements under which major digital companies **pay substantial remuneration to content providers** that is difficult for their competitors to replicate.
- Misuse of **GDPR** rules restricting access to data.
- Use of publishers' content by foundation model providers, **without the authorisation of rights holders**.



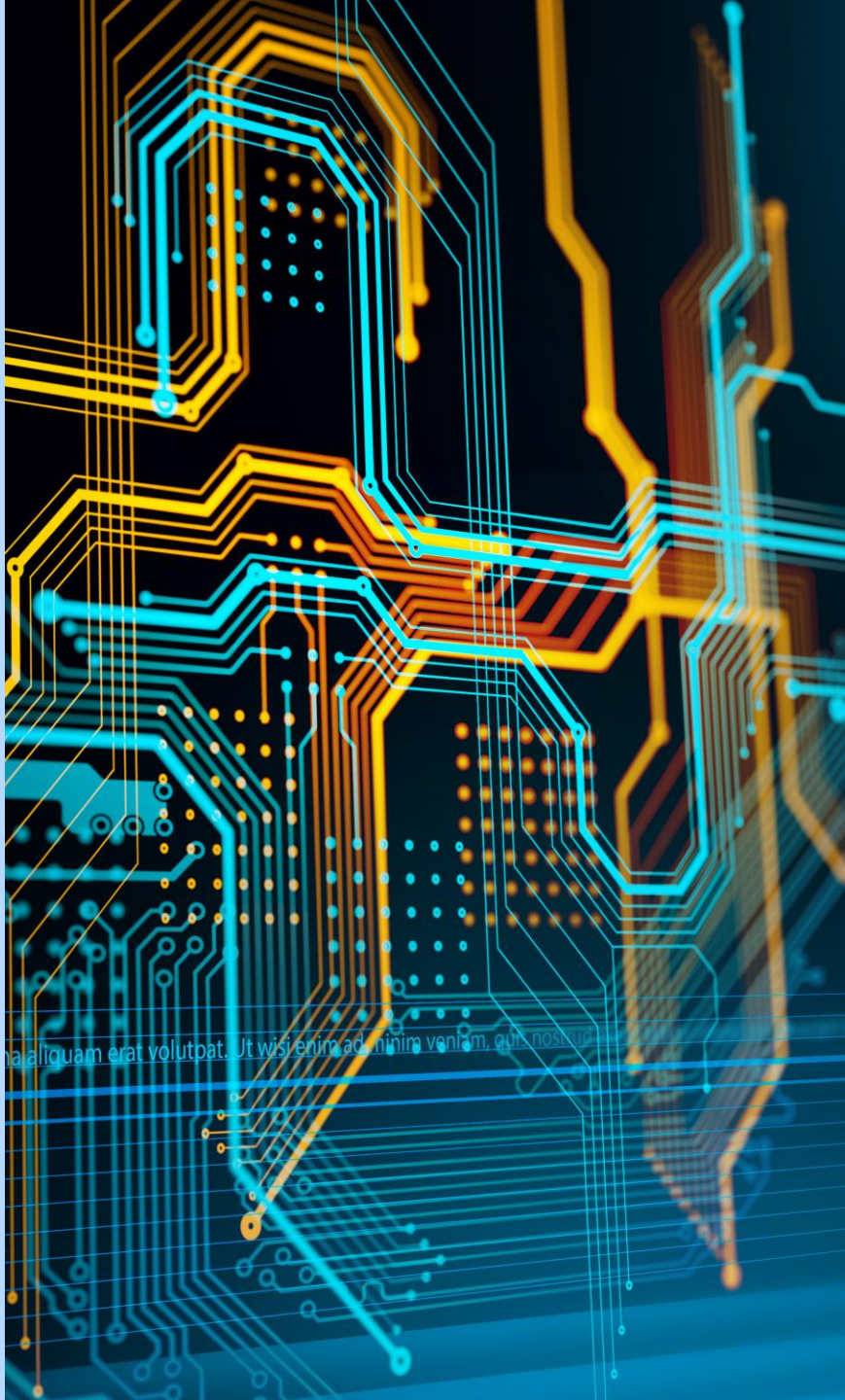
# THE COMPETITION RISKS (2)

- **Risks related to access to a skilled workforce:**
  - **No-poach agreements.**
  - **Recruitment by digital giants of strategic employees** (e.g. Microsoft's hiring of most of Inflection's 70-person staff) is also raising concerns.

- **Risks associated with open-access models:** the conditions of access and reuse of models or some of their components can lead to users being lock-in (e.g. "Google Android").

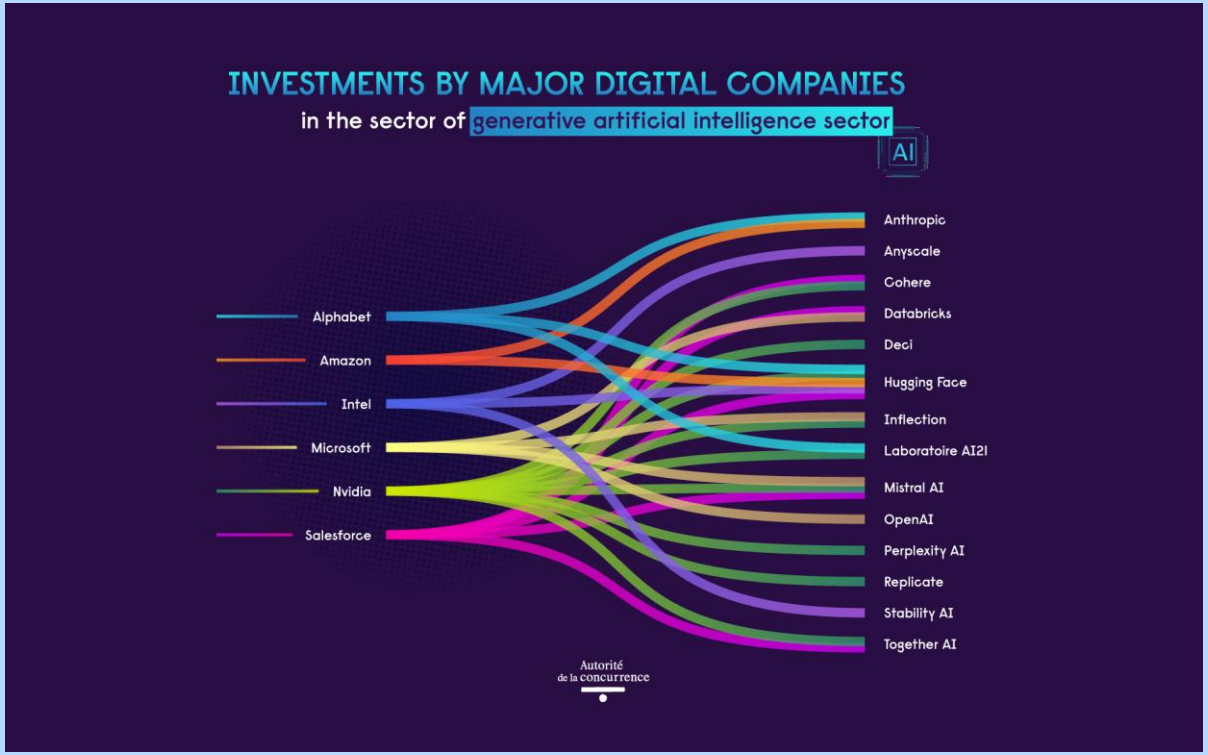
- **Risks associated with the presence of companies on several different markets:**
  - Upstream, risk of refusal of or limited access to the chips or data needed to train competing foundation models, as well as exclusivity agreements.
  - **Downstream, the integration of generative IA tools on certain devices, such as smartphones**, is raising concerns (e.g. the Google/Samsung agreement).

- Risk of concerted practices.



# THE COMPETITION RISKS (3)

- **Minority investments and partnerships by major digital companies may raise competition concerns.**
- **Need for vigilance given the competition risks:** weakening of competition, vertical effects, market transparency, lock-in of some parties, etc.
- These agreements may be examined by competition authorities on several legal grounds.
- However, the *Autorité* notes a **lack of transparency** in the agreements.



PART 3

# The recommendations made by the *Autorité*

# MAKE THE EXISTING REGULATORY FRAMEWORK MORE EFFECTIVE

Given the speed of change in the sector, existing regulations should be fully implemented and their impact evaluated, in order to avoid negative impacts on innovation and competition.

With no change to existing legislation, improvements could be made:

Proposal no. 1: the Commission should pay particular attention to the development of MaaS services to assess the possibility of designating companies providing such services as gatekeepers under the DMA.

Proposal no. 2: at the French level, in implementing the provisions of the SREN Law on cloud credits, the DGCCRF should pay particular attention to the use of such credits in AI.

Proposal no. 3: the AI Office, established under Article 64 of the AI Act, and the competent national authority in France, which will be designated in accordance with Article 70 of the AI Act, should ensure, on the one hand, that the implementation of the Act does not hinder the emergence or expansion of smaller operators and, on the other hand, that the largest operators in the sector do not misuse the text to their advantage.

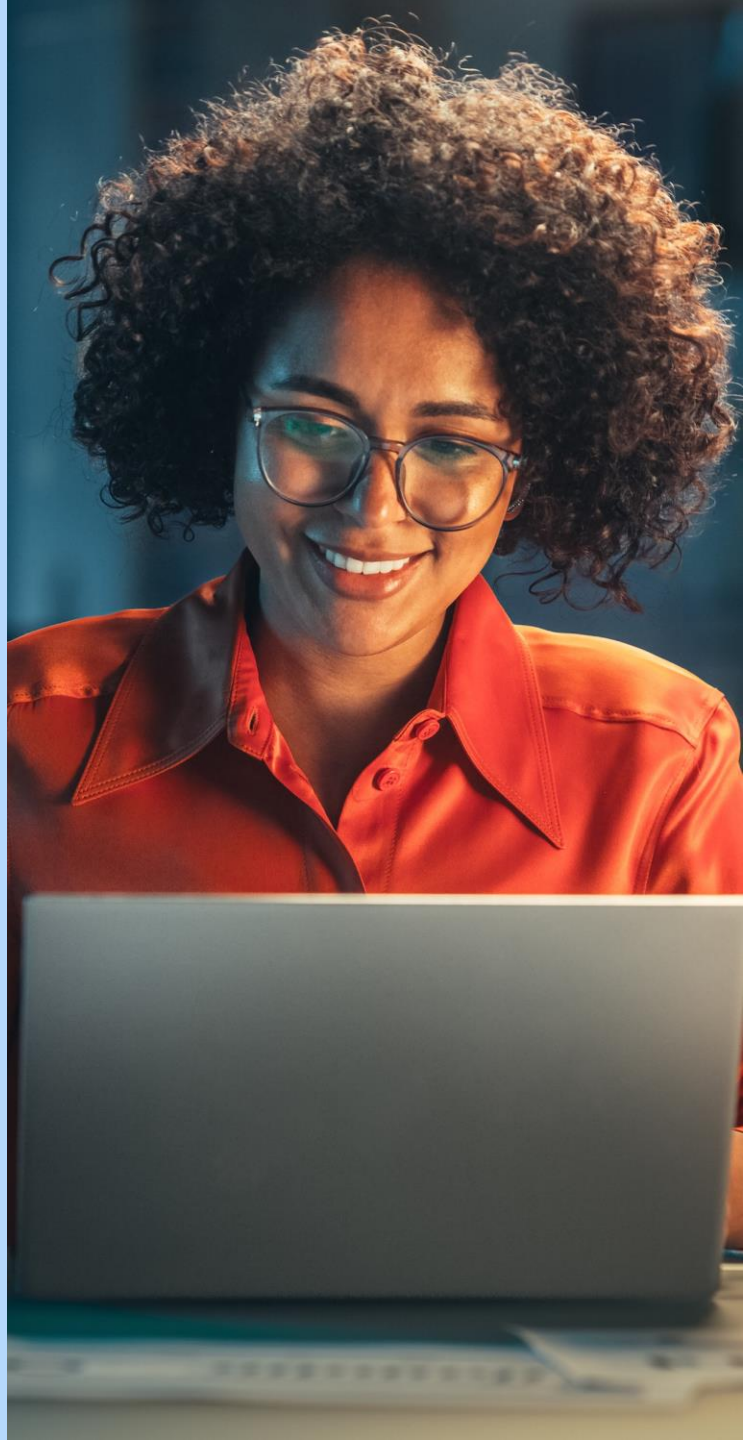
# USE THE FULL EXTENT OF COMPETITION LAW TOOLS

The use of competition tools, including with regard to restrictive practices that are primarily the responsibility of the DGCCRF, will play an essential role in preventing the emergence or consolidation of dominant positions or agreements that would affect competition in the sector.

The *Autorité* can draw on its experience in the digital sector:

- Joint study with the German *Bundeskartellamt* on algorithms (2019).
- Market study on competition in the cloud sector (2023).
- Unannounced inspections in the graphics card sector (2023).
- Google “related rights” case (2024).

Proposal no. 4: the authorities responsible for enforcing competition in the markets must remain vigilant in the generative AI sector and, if necessary, use all the tools at their disposal to act swiftly and effectively.



# INCREASE ACCESS TO COMPUTING POWER

Access to computing power is essential for the development of research and the emergence of new companies in the generative AI sector.

Proposal no. 5: continue to invest in the development of supercomputers at European level, to give as many parties as possible access to computing power.

Proposal no. 6: the government and/or companies responsible for managing supercomputers could look into how to propose 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.

Proposal no. 7: in conjunction with the AI Act in particular, set criteria for the openness of generative AI models trained on public supercomputers.

# TAKE ACCOUNT OF THE ECONOMIC VALUE OF DATA

At each stage of the value chain, data is used differently (training, fine-tuning or inference). Agreements between rights holders and developers should reflect the relative importance of the data for the developers. Attention should also be paid to transaction costs.

Proposal no. 8: public authorities, in particular as part of the mission entrusted by the French Ministry of Culture to the French Higher Council for Literary and Artistic Property, could encourage rights holders to take into account the economic value of data according to the use case (e.g. by introducing differentiated pricing), and to propose bundled offers to reduce transaction costs, in order to safeguard the innovation capacities of model developers.

Proposal no. 9: make public and private data available for the training or fine-tuning of generative AI models, and encourage public and private initiatives to distribute French-language data, whether text, image or video.

# ENSURE GREATER TRANSPARENCY ON INVESTMENTS BY DIGITAL GIANTS

Pending a decision by the CJEU on Article 22 of the Merger Regulation, the existing legal framework can be used to examine most competition concerns regarding agreements between companies, whether *ex ante* through merger control rules or *ex post* through competition law (cartels or abuse of dominant position).

Even with no change to existing legislation, however, there should be greater transparency on minority, non-controlling investments in the sector.

Proposal no. 10: the Commission could request further information on minority investments in the same sector as the target, in the template relating to the obligation to inform about a concentration pursuant to Article 14 of the DMA.