

CRITEO's Submission: European Commission's consultation on competition in Generative AI *March 2024*

Please find below CRITEO's contribution to the European Commission's Directorate General for Competition (DG COMP) consultation on competition in generative AI.

As a leading player in the AI sector, CRITEO has been at the forefront of the Generative AI revolution. Generative AI, and General Purpose AI (GPAI) more broadly, bear formidable opportunities both for fundamental research and for businesses. This revolution stems from decades of breakthroughs in the fields of computing power and algorithmic science, and will have a long-lasting impact on digital markets. It is therefore crucial to ensure that **all market players can seize the opportunities offered by this revolution** and that it **does not further entrench existing dominant positions or increase market concentration**. Ensuring the availability of AI models for businesses in downstream value chains (in particular open-source models), mitigating the risk of lock-in strategies in the cloud and ensuring a fair sharing of value between GPAI developers and users will be decisive in this regard.

While generative AI can be seen as a major technological breakthrough, **it may lead to competition issues that have already been identified in other technologies, but does not seem to give rise to new challenges from a policy and legal perspective**. Generative AI models share the same characteristics and dynamics than other digital services, with network effects translating into feedback loops for instance. In other words, the emergence of generative AI systems does not require an overhaul of competition rules or principles, but rather stronger enforcement by competition authorities. Indeed, generative AI has the real potential to become the "horror gallery" of antitrust law and dominance, with tying, lock-in, self-preferencing practices building higher barriers for new market entrants. CRITEO therefore welcomes the Commission's initiative to delve further into this market.

1. Preliminary remarks: How CRITEO harnesses the potential of AI models

As an AI-born company, CRITEO has of course been at the forefront of the foundation models/general purpose AI/ generative AI revolution. We have been among the first players in the advertising sector to use large general purpose AI models to further develop different AI use cases on the open Internet for instance as follows:

- As a technological block in CRITEO's AI supply chain to generate mathematical vectors or other kind of non human interpretable content, then reused in to feed other models;
- As a tool to improve CRITEO's AI led products for instance to generate ad creatives ;
- As a productivity tool to improve CRITEO's internal operational efficiency, for instance to develop commercial chatbots to interact with consumers, or to identify information across CRITEO internal knowledge base.

CRITEO is using both open-source and proprietary models, most of the time in combination with third-party solutions or our own open-source models.

2. Responses to the questionnaire

Please note that the following remarks relate to **general-purpose AI models** (hereafter “GPAI”), a more encompassing term which designate a type of AI that can understand, learn, and apply its intelligence to solve a wide range of problems without needing to be retrained for each new task. Generative AI is a subset of GPAI focused on the creation of new content based on the inputs it receives.

Questions

- 1) What are the main components (i.e., inputs) necessary to build, train, deploy and distribute generative AI systems? Please explain the importance of these components*
- 2) What are the main barriers to entry and expansion for the provision, distribution or integration of generative AI systems and/or components, including AI models? Please indicate to which components they relate.*
- 7) What is the role of data and what are its relevant characteristics for the provision of generative AI systems and/or components, including AI models?*
- 9) Do the vertically integrated companies, which provide several components along the value chain of generative AI systems (including user facing applications and plug-ins), enjoy an advantage compared to other companies? Please elaborate on your answer.*

More than for other AI systems, data is the most critical input to train and develop GPAI models, as massive amounts of data are required.

Access to data has become one of the most important barriers to entry and to growth in digital markets for different reasons :

- **A complex and inconsistent regulatory environment** : The delineation between personal data and non-personal data in EU law generates considerable legal uncertainty and disincentives businesses to share their data.

Indeed, the notion of personal data under the General Data Protection Regulation (GDPR) has been interpreted so broadly by Courts and national Data Protection Authorities (DPAs) that any kind of data can today qualify as personal data. The propensity of DPAs to consider that processing and worse sharing data are inherently risky, has raised strong barriers for the access of third-parties to personal data. The lack of harmonization in GDPR implementation across the EU also has an impact on EU players. These two factors undermine the objectives of the EU Data Strategy, which was intended to unlock the access to data to enable the training and development of AI models.

These regulatory developments should be better monitored by competition authorities as they have a significant impact on the competitiveness of EU players.

- **Anti-competitive practices by very large vertically integrated platforms** : The largest online platforms that have implemented various strategies to block or reduce access of third-parties to their users' data, adding new features or stopping technical support on key technologies, using potential risk to the protection of user personal data and risk to security (in most cases these risks are minimum) has a legal ground to actually foreclose competitors. Unfortunately, the Digital Markets Act (DMA) does not prevent such strategies.

The training of GPAI models also involves formidable **computing power**, that only the largest market incumbents in the cloud market are able to provide. The combination of their privileged access to first party and proprietary data for several years and their computing capabilities place them in an unrivalled position in the GPAI revolution.

Questions

3) What are the main drivers of competition (i.e., the elements that make a company a successful player) for the provision, distribution or integration of generative AI systems and/or components, including AI models?

4) Which competition issues will likely emerge for the provision, distribution or integration of generative AI systems and/or components, including AI models? Please indicate to which components they relate.

5) How will generative AI systems and/or components, including AI models likely be monetized, and which components will likely capture most of this monetization?

8) What is the role of interoperability in the provision of generative AI systems and/or components, including AI models? Is the lack of interoperability between components a risk to effective competition?

For most companies, it is more interesting to **fine-tune existing models** and to **integrate them into downstream systems or applications** than to develop new ones. The idea behind GPAI models, also known as foundation models, is precisely to perform a wide range of tasks by transferring knowledge across different domains. In a sense, it would defeat the purpose of foundation models to have every companies building their own.

Fine-tuning basically means tailoring a model to specific needs by taking a pre-trained model and further training it on a specific, often smaller, dataset to specialize its performance for particular tasks or domains. This process allows the model to adapt its general knowledge to the nuances and specifics of a targeted field or type of task, improving its accuracy and effectiveness for those specific applications.

In this context, CRITEO calls EU competition authorities to :

- **Ensure the availability of GPAI models for businesses in the downstream value chain :** Today, the leading GPAI models on the market are either developed or distributed by the largest online platforms. Considering that these same players are often vertically integrated, they may be incentivized to strategically limit access to their models to competitors in downstream or adjacent markets. Therefore, it is important to ensure that access to these models is granted under Fair, Reasonable and Non-Discriminatory (FRAND) conditions for all market participants.
- **Mitigate the risk of lock-in when AI models are distributed in the cloud :** Today, only the largest providers of cloud computing services have the capacity to distribute and enable the fine-tuning of the main proprietary models on the market. Cloud based models provide a ready to use/off the shelf solution enabling AI developers to iterate and develop quickly their first proof of concept. This raises competition issues since the cloud is captive by nature. Indeed, the input, the model and the output remain in the cloud, with no access to the model architecture and the model weights (i.e., the parameters of the AI models resulting from the training phase) and therefore no ability to export it. Barriers to data portability and interoperability could therefore raise lock-in effects.
- **Ensure a fair sharing of value between GPAI developers and users :** Like network effects for online platforms, there is a real risk that the main providers of proprietary models could take a massive advantage over competitors through feedback loops – basically the use of information generated by the use of their products by firms like Criteo to further improve their own models. This raise the crucial question of the sharing of value between developers of proprietary models and fine-tuners of such models. It is key for competition enforcers to ensure that the sharing of value in the conditions of use of these models is fair and do not benefit to the incumbent only.

Questions

6) Do open-source generative AI systems and/or components, including AI models compete effectively with proprietary AI generative systems and/or components? Please elaborate on your answer.

Companies should be able to find viable alternatives to the main proprietary models and cloud solutions. Facilitating open-source development and deployment in companies is a solution.

For companies like CRITEO, open-source models bear a lot of opportunities, both in commercial and scientific terms. Access to the codes and models' weights enable our researchers to improve and tailor them to our specific needs. Open-source models, in particular pre-trained models, significantly reduce barriers to entry as they enable smaller firms to innovate at substantially reduced costs.

It should be recalled that open-source is a very broad term, as there can be various degrees of openness and several ways to open a model. The underlying code, the architecture, the

training datasets, the weights or the training process can be opened together or separately. The use of open-source models is also framed by licensing terms and conditions, which sometimes excludes the commercial use of the model.

To have thriving and dynamic AI ecosystem in the EU, it is crucial for competition enforcers and policymakers to ensure the availability of open-source models.

Questions

11) Do you expect the emergence of generative AI systems and/or components, including AI models to trigger the need to adapt EU legal antitrust concepts?

12) Do you expect the emergence of generative AI systems to trigger the need to adapt EU antitrust investigation tools and practices?

The principle-based approach of EU antitrust law is flexible enough to deal with the challenges of the generative AI revolution. Current market distortions can be addressed through existing concepts (such as tying or self-preferencing) and procedures.

In parallel, it is necessary to further develop the ex-ante regulatory approach to prevent abusive or harmful practices before they occur. The Commission should notably consider qualifying GPAI models as a core platform service (CPS) under the DMA subject to non-referencing obligations and interoperability obligations. The scope of several DMA provisions, such as the obligation to give access to business users on FRAND terms could for instance be expanded to this new CPS.

Given that cloud services are critical for the training and distribution of AI models, the Commission should also consider designating the largest providers of these services as gatekeepers. The market for GPAI models would indeed benefit a lot from increased competition in the cloud market.

CRITEO's AI LAB : The first European private AI laboratory in terms of scientific publications

CRITEO is a native French AI company, which grew up as the first global adtech firm in the open Internet. Our success was built on a recommendation algorithm, which was first developed for recommending movies and then commercially used to recommend products based on previous interactions.

CRITEO has today one of the largest AI Lab in France and Europe, located in Paris. Per the number of publications in top-tier AI venues (conferences and journals), the Criteo AI Lab is the first private European AI laboratory.

CRITEO's research covers the whole AI spectrum, including privacy-enhancing technologies, recommendation systems, generative AI and deep learning for text/image/video, causality and fairness. Since 2018, the CRITEO AI Lab published around 220 research papers including 25+ research papers on generative AI.

Since our creation, we've been strongly committed to sharing value across the entire ecosystem: All our research papers are in open access and we are also putting open-source code and anonymized open-source datasets at the disposal of the research community.