COMPETITION IN GENERATIVE AI – EUROPEAN COMMISSION'S CALL FOR CONTRIBUTION – OVHCLOUD'S CONTRIBUTION

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

Building, training, deploying, and distributing any AI system, including generative AI systems requires 4 main types of components or assets that are of equal importance:

- <u>Infrastructures</u>: AI, and particularly training AI models, requires significant computational resources, which are necessary to process the data, train the models and then deploy it. This compute power translates into dedicated hardware with specialized chips, mainly graphical processing units (GPUs) which are able to provide different levels of compute power depending on the user's needs, on demand. Such resources can be accessed directly through the purchase of a GPU, although this may be expensive and complex for a small or medium-sized company, or more easily through cloud providers giving access to such GPUs through their cloud services.

- <u>Data</u>: essential for the training or fine tuning of any AI model. As recently <u>pointed out</u> by the United States FTC, developing AI and specifically generative AI requires exceptionally large datasets, especially in the pre-training steps. Other than the amount of data, its quality is also crucial for the effective training or fine-tuning of any AI model, granting a significant competitive advantage to any company benefitting from a privileged access to such data. This includes both 'public' data and 'private' data (organizations' own data). For the latter, given the importance and sensitivity of such data, ensuring they are exploited within a secured environment, fully compliant with EU legislation and preventing any unlawful access by third parties, will be key to maintain privacy data protection in the age of AI and therefore ensure a trustworthy innovation.

- <u>AI platforms and solutions</u> bringing together the necessary tools (e.g. frameworks, algorithms, machine learning operations, datasets transformation etc.) necessary for the exploitation of data by a model, and for the training or fine-tuning of such model. Being able to provide AI solutions that are accessible and suitable for all types of organizations (including SMEs) would improve their adoption and, as such, contribute to accelerate the digitalization of the EU economy.

- <u>Human expertise / Competencies</u>: developing an AI model requires very qualified talents – especially in engineering and research – with a specific set of skills in machine learning, natural language processing, and computer vision. This is particularly important in a rapidly evolving technology such as AI, where any model or solution can quickly become obsolete. Benefitting from the right talents to maintain relevance in this rapid evolving technology is thus key.

OVHcloud

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.

Several barriers to entry and expansion can be identified for the provision, distribution or integration of generative AI systems and/or components. Below are the ones considered to have the most significant impact on the market:

- Access to the necessary infrastructures: as explained in answer to Q1, the training of AI models relies heavily on GPU and GPU-like processors, which can be a significant investment for any organization. The cost for high-performance GPUs for instance can be estimated between €10,000 to €25,000 each, an amount that many companies cannot afford. Given the very high demand for such GPUs that the emergence of generative AI has led to, accessing the necessary infrastructure in a timely manner represents a great challenge both for companies wishing to develop and train their model, but also for cloud providers wishing to offer such GPUs through their services. Not benefitting from a secure access to these resources can thus represent a major barrier to entry and expansion within the AI market for such organizations.

- Access to data: different AI models/solutions providers might not have the same access to data depending on their other/adjacent activities. The volume and quality of data required to pre-train a generative AI model from scratch may impact the ability of new players to enter the market. As pointed out by the US' FTC, more established companies may benefit from access to data collected from their users over many years—especially if the incumbents also own digital platforms that amass large amounts of data. Such companies are also more likely to have developed data collection tools and technologies for acquiring or scraping data. This grants them a significant competitive advantage in creating high quality and performant AI models, given the structural role data plays in it. On the other hand, not having access to important amounts of quality data can thus constitute an important barrier to entry and expansion for the provision of AI models.

- Access to technical expertise: as pointed out in the FTC conclusions, labor expertise is an essential input for generative AI, as generative models require significant engineering and research workforce with particular skillsets as well as a deep understanding of machine learning, natural language processing and computer vision. Thus, difficulties to access to such talents can represent a major barrier to entry and expansion within the AI market. Having said that, it is important to note that not every organization needs to benefit from the most skilled/expert employees, as this mainly depends on the usage / AI application it chooses to implement within such organization. While certain organizations might need very expert data scientists to develop their own specific AI application, other will be able to simply rely on developers to adapt existing applications/models to the organization's needs. It is therefore of relevance to develop AI solutions that are accessible and suitable for all types of organizations (including SMEs).

OVHcloud

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?

The main drivers of competition for the provision, distribution, or integration of generative AI systems and/or components are directly related to the main barriers to entry mentioned in answer to Q2. They are:

- Access to the necessary infrastructures: as mentioned in answer to Q1 and Q2, GPUs are essentials for the training and deployment of any AI model. Given their high prices and the current massive demand, being able to offer such GPUs is key for any company wishing to position itself in the AI market. A privileged access to such infrastructures would undoubtedly confer a significant competitive advantage to any company, particularly any cloud provider, as its services would become very attractive for any company wishing to develop or implement AI.

- Access to data: data being at the foundation of any AI model, the quality and quantity of data a company can use to train its AI model will have a direct impact on the performance of the end model. This has the potential to favor large platforms benefitting from strong databases issued from their activities in adjacent markets.

- Competencies: the number of AI experts, able to master the development of AI solutions and models, being at this stage limited confers a significative advantage for companies benefitting from such talents within their workforce and allows them to develop more innovative solutions.

Other drivers of competition can be mentioned:

- Data protection / sovereignty: organizations wishing to implement AI applications based on their own 'private' data will require their provider to ensure such data, which might be of very sensitive nature, is exploited within a highly secure environment and fully compliant with EU legislation. As such, providers offering this level of privacy and data protection, including against third-country access enabled by certain extraterritorial laws, will benefit from a competitive advantage within the market.

- Accessibility: AI components, especially infrastructure (i.e. GPUs), being of very high cost, they lay the risk to see smaller organizations (e.g. SMEs) not being able to afford them and thus to implement AI within their process. In this context, providers able to offer such solutions at an accessible price will be very competitive for this market.

- Reputation: whether a company's AI solutions are considered as performant and communicated as such to the public can have strong impact on steering users towards such solutions. In that sense, the public support expressed towards a company, whether it is by users, partners, global companies, or public authorities themselves can greatly contribute to the reputation of a company and increase its ability to compete within the market.

- Distribution: distribution agreements either with model providers or cloud providers can impact both type of companies' ability to attract customers. For instance, an AI model provider who does not have distribution agreements with the main cloud providers, or repository providers, will face difficulty in making its models available to the end AI users,

and thus will not be competitive. On the opposite, a cloud provider or repository provider who does not have distribution agreements with the main AI models providers will end up with a limited AI offer which might steer AI users towards the services of other providers.

- <u>Sustainability</u>: with AI relying on very high compute power, thus very energy consuming, maintaining a low carbon footprint will be a challenge for any company wishing to implement the technology. Offering AI resources based on sustainable services, could thus represent a competitive advantage for any AI provider, as it would allow its users to implement AI without jeopardizing its sustainable commitments.

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.**

Many competition issues are likely to emerge, or have already emerged, for the provision, distribution, or integration of generative AI systems and/or components. Among them, the main ones are:

- <u>Financial lock-in of users</u>: with cloud services providing access to GPUs, competition will raise among cloud providers to capture the AI market wanting to use such GPUs for their AI development, and the associated growth issues from this market. However, so-called cloud "hyperscalers" are currently using their superior means, some of them issued from activities unrelated to cloud services, to offer cloud credit (i.e. free access to cloud services) with the objective to capture the AI market within their services. The "hyperscalers", who already had Startup programs providing cloud credits <u>up to</u> 200,000€ for companies, have recently increased the amount of cloud credits they offer to AI companies (<u>up to</u> $350K depending on the provider) to an amount that alternative providers simply cannot match. Given the fact that AI model's training or fine tuning is particularly expensive, due to the necessary investments in compute capacities, this has for effect to encourage AI users to choose the services of these "hyperscalers", not because they would best fit their needs in the long-term, but because of the amount of cloud credits which has been offered to them. As such, these cloud credits programs constitute a way for "hyperscalers" to unfairly capture the AI market and lock it into their services.

- <u>Technical lock in of users</u>: cloud "hyperscalers" are offering proprietary solutions (e.g. Automated Machine Learning, fine tuning), for users wishing to create or fine-tune their models more easily. However, when the end model is eventually created, users do not get access to the model itself (i.e. its file) but solely have the possibility to use it or to deploy it from the provider's infrastructures. This has for effect to lock the user into the service of the provider offering the training solution as (i) given that the training is done by the provider itself, with no transparency, the user does not have an understanding of his model's functioning (e.g. with which algorithm it has been tested) (ii) the training solution is a proprietary one, not open source, which significantly limits the user's capacity to migrate or recreate his model in another provider's infrastructure (iii) the user does not have access to his model (i.e. the file), and thus cannot migrate it to another provider. In the end, the user is effectively locked in the services of the training provider, as he would have to recreate his model almost from scratch should he want to use it in the services of another provider. This greatly limits alternative providers' ability to compete for the AI market and has the

potential to reinforce the cloud market's consolidation around the "hyperscalers", as the model trained on their services can only be exploited from their cloud infrastructures.

- Access to data: as described in answers to Q1 and Q2, the role of data is essential in the development of AI models. Large companies and platforms benefitting from privileged access to quality data issued from their adjacent/historic activities thus enjoy a substantial competitive advantage in the development of performant models. The lack of requirements for such companies to be transparent and explain on which type of data / datasets their models have been trained might allow them to increase their competitive advantage. This is because, as such, alternative companies will have no knowledge about the way such models are developed and will not be able to try and recreate alternative ones, thus conferring a substantial competitive advantage to the company benefitting from the privileged access to data.

- Tied selling: companies having activities in markets adjacent to AI, particularly the software market, might use their position in these to unfairly promote their own AI solutions. This can be the case through the tying of different products between each other, for instance by integrating an AI solution directly into a software suite or offer. This has the potential to significantly consolidate the AI market around already dominant digital companies, by strongly encouraging software users to use the AI products developed by their software provider, and eventually becoming dependent on such products.

- Privileged access to infrastructures: as mentioned in answer to Q1 and Q2, the access to sufficient infrastructures to train and deploy AI models is key in a company's ability to position itself in the AI market. Given the massive current demand for GPUs, access to this infrastructure is complex for any cloud provider, thus conferring a strategic advantage to any company benefitting from a privileged access to GPUs. Public authorities should ensure that strategic partnerships between companies do not lead certain providers from enjoying an unfair access to these infrastructures at the expense of other providers, as this could constitute a substantial competitive distortion within the market.

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

Given the current market structure and demand for GPUs to train and deploy AI models, it is likely that this component will capture most of the market's monetization in the short term.

However, as the GPU market will become more competitive, most of AI's monetization is likely to be captured by companies offering AI applications and companies offering consulting and integration of AI inside companies' internal systems.

There are two main segments of customers for generative AI systems: consumers and professionals. While the consumer market may see significant disruption as AI becomes more integrated into daily life, the professional market is likely to be the one to issue most revenue from AI given the advantage the technology represents in cost savings, improved quality, and optimization of business processes.

As such, it is likely that most of the monetization for generative AI systems will come from applications that are integrated into the internal systems of businesses, in the same way enterprise

software have captured most of the software market's monetization. This can include the addition of AI capabilities to existing enterprise software, such as automated CRM tools, customer support, productivity suites.

Two other components are also likely to capture significant revenues from generative AI monetization:
- Cloud services: which will be at the core of any AI model and/or application and will thus capture a substantial part of the associated revenues.
- Consulting and integration services: which will often prove to be a necessity for companies wishing to integrate AI within their internal systems.

As such, verticalized companies, for instance offering both AI software for the professional market, as well as cloud services, are likely to capture a significant part of generative AI's monetization. This calls for a particular scrutiny from public authorities to prevent such companies from engaging in unfair competitive practices within the market, for instance to favor the use of their AI application / software on the basis of their cloud services and at the costs of consumers' choice.

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.**

There are examples of models developed by companies which show that open-source AI models can offer the same range of performance as proprietary AI models, and thus compete effectively with them. It is important to note that the development of such performant models is made possible by the contributions received from the open source community, for instance in recent years with the release of large foundation models in open-source. Such releases allow the AI ecosystem to significantly improve its knowledge/expertise of the technology and to build on top of it.

Leaving proprietary solutions prosper without setting rules for fair competition might have the following negative effects:

- Lock in of users: as explained in answer to Q4, proprietary training or fine-tuning solutions have for effect to lock the user and its model within the infrastructures of its training provider, thus having adverse effect on competition.

- Concentration of AI open source within large companies given the fact that the development of foundation models requires massive investments, especially in computing capacities, it might encourage companies developing such models to offer them in proprietary formats, to ensure their profitability.

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?**

As pointed out by the US FTC, data are at the foundation of any AI model. The training of such model requires exceptionally large amounts of data, with the quality and quantity of such data having a direct impact on the model's performance and the trust the end user can have in it (more qualitative data means more accuracy and performance, and less bias).

Due to this crucial importance of data, large platforms with massive amounts of qualitative data originating from their customers' historic use of their other services enjoy a substantial competitive advantage. Indeed, their access to large amounts of quality data gives them a great advantage in the training of powerful and qualitative models, which their competitors might not be able to match due to the lack of similar access to data.

While forcing such companies to disclose the data used for the training of their model might be difficult due to business secrecy or data protection, requiring them to communicate on the type of data which have been used would be a first step for both (i) AI models' transparency, for instance to help identify potential bias and correct them (ii) more competition, as AI companies wishing to compete with these providers but not benefitting from the same access to data would be able to understand the way these models have been designed/trained and thus to offer alternative offers. On the other hand, and as mentioned in answers to Q1 and Q3, 'private' data will also be a key component in the development of the AI market, as organizations will want to implement AI applications/models based on their own specific data for improved efficiency and to answer their specific needs. Given the high sensitivity of such data, implementing strong data protection requirements along the development of AI will be key to protect organizations' data, including their most strategic ones (i.e. intellectual property rights).

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?**

As explained in answer to Q4, offering certain solutions (AutoML, fine-tuning) only in a proprietary format without adequate legal framework can lead to users' lock-in within the infrastructures of the training provider. As such, models trained from these solutions can only be used within the environment of the training provider, thus preventing the user from migrating it to another provider, even though the latter might best fit the users' evolving needs. This directly impacts users' freedom of choice, and as it has been the case for other markets (e.g. cloud), calls for more interoperability, portability, and reversibility requirements throughout the development of the technology, to prevent user lock in.

Another negative effect of this lack of interoperability is the fact that such proprietary solutions lead the model to being exploited only within the cloud infrastructures of its training providers. With such training solutions being currently offered by the major cloud "hyperscalers", this has the potential to significantly reinforce the cloud market's consolidation around such players over the course of AI's adoption.

On the other hand, interoperability between components and open-source in general have proven to have a positive impact on the adoption of AI models. An example of this is the number of AI repositories which are based on open-source requirements. These allow the user to access AI models and use different ones in parallel or combination in a much easier way, thus favoring the use and adoption of the technology.

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.

The vertically integrated companies, providing several components along the value chain of generative AI systems may enjoy a substantial competitive advantage to other companies.

The following examples of verticalizations offer the biggest competitive risks:

- Companies offering both cloud services and manufacturing GPUs: given the very high demand for GPUs and the ongoing competition between cloud providers to offer them through their services, a company vertically integrated and providing both these components may have a substantial competitive advantage over the AI and cloud markets, depending on its size and global capacities. The risk would be to see such companies being in competition with its own customers (as it would be providing GPUs to other cloud providers) and providing an unfair access to its GPUs to its own cloud services.

- Companies offering both AI models and benefiting from a privileged access to data: as explained in previous answers, companies with such verticalized activities, especially when also operating historic activities in adjacent markets, benefit from a strong competitive advantage in the provision of AI models. Their access to large amounts of quality data, enables them to provide for more performant models, that companies which do not operate activities in the data industry might not be able to match.

- Companies' dominant in the software market and offering AI solutions: such companies might enjoy a substantial competitive advantage in the adoption of their AI models, particularly as they could engage in anticompetitive practices (e.g. tied selling, bundling, self-preferencing) to unfairly promote their AI solutions.

- Companies offering GPUs and the software to exploit this infrastructure: with the current market structure of the GPU market, a company having control over both the infrastructure itself, and software programs to exploit such GPUs, including software allowing a company to offer AI models for inference, would gain great control over the AI market. Cloud providers, for instance, would end up in great dependency towards such GPU provider, as it would depend on this provider granting it access to its solutions to exploit the infrastructure (incl. for inference). Should this not be the case, the cloud provider would lose the ability to offer the maximal exploitation of the GPU, and to offer AI models for inference, and thus have a significant competitive disadvantage compared to its competitors.

10. What is the rationale of the investments and/or acquisitions of large companies in small providers of generative AI systems and/or components, including AI models? How will they affect competition?

While OVHcloud cannot comment on the rationale of such investments or acquisition, it can offer insights about the way they would affect competition. The main estimated impacts could be, in coherence with answer to Q9:

- Investments and/or acquisitions between companies manufacturing GPUs and providing cloud services: this would lead the GPU provider to compete with its own customers and

strongly encourage him to grant privileged access to its infrastructures to the cloud provider it has ties with. In an AI market, and more generally cloud market, where access to such infrastructure is essential and complex, such situation would create a critical distortion of competition and would grant the cloud provider with a significant and unfair competitive advantage.

- Investment and/or acquisition between a dominant software company and AI models: such situation could lead to converting the dominant position in the software market into a dominant one in the AI market, by tying, bundling or self-preferencing practices.

- Investment and/or acquisition between a company benefiting from a privileged access to data and a company providing AI models: such partnership would allow the companies to develop superior AI models due to the privileged access to large amounts of qualitative data other AI models providers might not enjoy, thus leading to an unfair advantage to such company.

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?

AI provides great prospects for the European economy, but also leads to new competitive risks within digital markets. There is a major risk that the emergence of AI contributes to reinforcing already dominant positions within such markets, particularly in the ones adjacent to the AI market (e.g. cloud, software). As such, the EU needs to have a clear understanding of the competitive risks linked to the emergence of AI, so as to adjust its current competitive enforcement tools to these dynamics.

Given the rapid pace at which the technology and associated unfair practices are developing and taking into account the historic "winner takes all" trend which has ruled digital markets, any regulation of the competitive risks linked to AI should be *ex ante* rather than *ex post*. This would ensure the EU corrects these unfair practices before they enable the reinforcement of current dominant positions.

To this end, OVHcloud is very much aligned with the proposal made by the European Parliament in its annual report on competition, to have the European Commission launch a market investigation under Article 19 of the Digital Markets Act, so as to evaluate the possibility to include generative artificial intelligence services of the gatekeepers under its obligations. The launch of this market investigation would lead the European commission both to:

- grasp a clearer understanding of the current competitive dynamics within the generative AI market and its impact on digital markets as a whole.
- have the possibility to adjust the DMA to match these competitive dynamics and prevent anticompetitive practices linked to the emergence of generative AI.

Depending on the conclusions of such investigation, *ad hoc* regulation might be needed to correct anticompetitive practices currently not covered by the DMA, as it has been the case in the cloud sector through the recently adopted Data Act.

OVHcloud

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

As described in answer to Q4, many competition issues are likely to emerge and in consequence, the European Commission and competition authorities will have to continue to examine and monitor issues resulting from generative AI.

OVHcloud notes that there already exist a European antitrust regulatory framework including merger control tools. The DMA will also reinforce the vigilance of the competition authorities by imposing enhanced obligations on gatekeepers, including a requirement of transparency on their possible acquisitions.

However, many anti-competitive practices may not be addressed by current EU rules. OVHcloud is concerned that the practices observed in the cloud market are being repeated on the AI market.

Those practices have been described along this consultation: tied selling, financial or technical lock-in, self-preferencing, unjustified lack of interoperability and portability etc.

With regard to all these elements, OVHcloud believes new tools might be needed to prevent this kind of anti-competitive practices. New regulations might also be useful as it is not yet possible to assess the impact of these kind of practice on the market and the direct or indirect harm resulting from it in the medium or long term.