

Competition in Generative AI

A Contribution to the EU Commission Call 2024

Prof. Dr.-Ing. Stefan Wagenpfeil
PFH Private University of Applied Sciences
E-Mail: sw@stefan-wagenpfeil.de
Phone: +49/2351/8959895

To: COMP-GENERATIVE-AI@ec.europa.eu

1. What are the main components necessary to build, train, deploy and distribute generative AI systems? Please explain the importance of these components.

Building, training, deploying, and distributing GenAI systems involves a number of key components playing an important role in the development and utilization of such systems. 1) Data is the foundation of GenAI. High-quality and diverse datasets are essential for the training of models. 2) The choice of the model architecture determines the model's capacity to learn and generate. 3) Training GenAI models is computationally intensive and time consuming. Proper infrastructure is necessary for effective training. 4) Pre-training on a large dataset and fine-tuning on specific tasks help models to generalize and adapt to various applications. 5) Ensuring the ethical use of AI, preventing misuse, and addressing potential biases is critical to maintaining trust and avoiding harmful consequences. 6) Deploying GenAI requires a robust and scalable infrastructure. 7) The way users interact with GenAI systems significantly impacts their usability and the system's acceptance. 8) Continuous monitoring of the model's performance and behavior after the deployment. 9) Adhering to legal and regulatory requirements is essential to avoid legal issues and reputational damage. 10) Deciding how and where to distribute GenAI systems including the decision if APIs are provided or how the system should / can be integrated into other applications. 11) Providing clear documentation and education resources for developers and users. 12) Ensuring the security of the model and user data in terms of privacy, data protection, denial of service, or system vulnerabilities. 13) Addressing bias in GenAI systems.

2. What are the main barriers to entry and expansion for the provision, distribution or integration of GenAI systems and/or components, including AI models? Please indicate to which components they relate.

Of course, each of the components described above provides a barrier as such. E.g., accessing enough and validated data is a barrier, providing the infrastructure, as well. However, some of these barriers should be highlighted as they are harder to overcome than others. For example the infrastructure for training: Meta recently announced to buy 350.000 NVIDIA A100 cards (each ca. 18.000 €) for the training of its own model. This means, to be competitive in this area, a company must invest more than 6 billion €. There are not many companies in the world able to do this. The second important barrier is access to training data (preferably in a legal way). To train such models trillions of records

are required. From my perspective, training data and the infrastructure are the most important barriers.

3. What are the main drivers of competition for the provision, distribution or integration of GenAI systems and/or components, including AI models?

There are two major important drivers for the competition: the access to data and the speed of development. That's the reason why a lot of money is spent to build training infrastructure, as this directly reduces the time-to-market of GenAI models. However, the second important driver (access to data) cannot be resolved by buying infrastructure. It only can be resolved by getting (or purchasing) the permission to access training data from social media, newspapers, search engines, etc. Therefore, the big players in the market (Google, Meta) have a huge advantage, as they already have existing databases usable for the training of GenAI models. Furthermore, the model architecture and its corresponding capabilities has a huge impact on the competition, as it defines the behavior and possible business cases of later applications.

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

The GenAI market is dominated by big companies. Monopoly and dominance are likely to exist making it difficult for smaller players to compete. The access to and ownership of data is a huge driver of competition, as well as disputes about intellectual properties. Failures to address ethical concerns and bias in AI models can result in public backlash and damage a company's reputation. Stricter regulation can create barriers and/or force market participants to pay penalties or to expensively rework their solutions. Mishandling of user data or privacy breaches can lead to legal challenges, reputational damage, and a loss of trust.

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

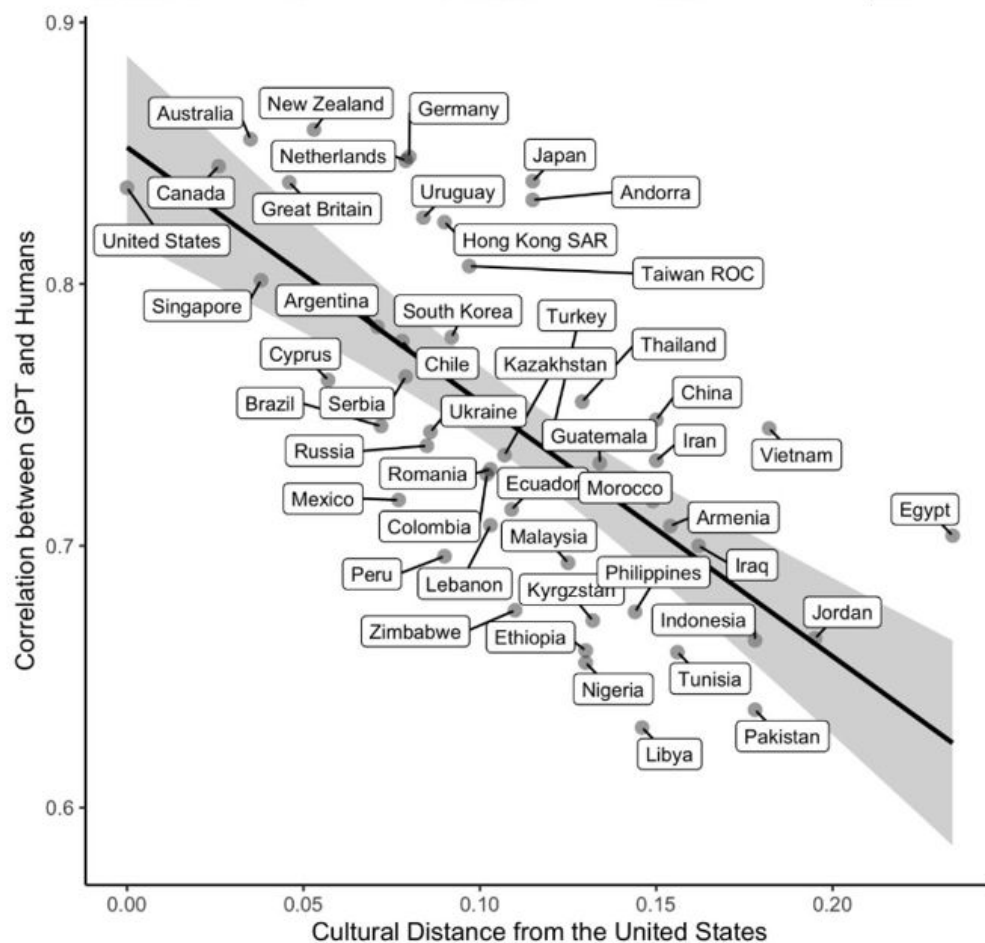
Monetizing GenAI systems mostly follows the typical monetizing models of other applications: subscription models, pay-as-you-go, licensing and royalties, data services, customizing and consulting, API access, premium features, advertising, education and training, marketplaces and app stores, research collaborations. From my perspective, these models are not particularly special in the area of GenAI. It has to be distinguished between the GenAI providers (like, e.g. Open.AI, Meta, Google) and the companies building services upon the provided models. A good example is Microsoft's Copilot service, which actually uses Open.AI's GPTs. While GenAI providers will look for subscription models and licensing, service providers might utilize the other forms of monetization more frequently.

6. Do open-source GenAI systems and/or components, including AI models compete effectively with proprietary AI generative systems and/or components?

From my perspective, open-source GenAI systems cannot compete effectively due to the missing hardware-infrastructure and – in many cases – due to the missing source data. However, open-source models can help to provide businesses based on these models.

7. What is the role of data and what are its relevant characteristics for the provision of GenAI systems and/or components, including AI models?

The role of data is one of the most important things in GenAI. Models, model architectures, services, can only be as good as the underlying data. And “good” is already a point that needs to be investigated. Depending on the outcome of GenAI, the term “good” also includes social, cultural, ethical, or political values, that can differ a lot throughout the world. One example is the cultural distance detected in Chat GPT, which shows that this model has been streamlined and mostly reflects western cultures (see following image).



Such effects result starting with the selection of input data. Furthermore, there are many examples, where input data has been used to train models without proper permissions. For example, the New York Times sued Open.AI and Microsoft for copyright infringement, as many articles from the New York Times' protected areas seem to have been included in Open.AI's training data. As GenAI is not explainable (or mathematically proof), it is hard to identify and/or validate such data misuse. Therefore, it would be good to understand, which data from which sources has been used for the training of models. The input data needs to be well balanced, compliant to laws and ethics, large enough to provide reliable results. All in all, a huge challenge.

8. What is the role of interoperability in the provision of GenAI systems? Is the lack of interoperability between components a risk to effective competition?

Interoperability needs to be established on a technical / service level. This means, that GenAI solutions need to provide their service in standardized technical formats (e.g., Webservices, REST-Services, JSON) in order to receive commands from the underlying models and construct automatized business processes and applications. Users will enter commands in most cases in their natural language, and the models will “understand” what results are expected. However, to achieve automation, the models have to be able to actively call 3rd party systems, which is typically done via API and the above-mentioned technical formats. An interoperability between models can be achieved by using natural language as a common level for exchange.

9. Do the vertically integrated companies, which provide several components along the value chain of Gen AI systems (including user facing applications and plug-ins) enjoy an advantage compared to other companies?

Similar to App-Stores, platforms from the big GenAI players (e.g., Open.AI, Google, Meta) provide development opportunities to everyone. Individuals, startups, existing companies of any size can build products or whole businesses based on these platforms. With extremely reduced development cost (as models are pre-trained), and a pre-defined commercial system (GenAI-App-Store, subscriptions, in-app-purchases), lots of efforts can be reduced. However, such vertically integrated companies are tightly embedded into the platform’s ecosystem and highly dependent on the quality and featureset of the underlying model. From a development and time-to-market perspective, this is a clear advantage. However, from a strategic perspective, this dependency has to be considered.

10. What is the rationale of the investments and/or acquisitions of large companies in small providers of GenAI systems? How will they affect competition?

At some point of time, huge investments in hardware are necessary when building GenAI models. Small companies will not be able to afford this. However, small companies might be able to introduce innovative ideas and model architectures. At that point, large companies might acquire smaller companies in order to provide their (existing) hardware and utilize the innovative ideas of smaller companies. In any case, GenAI will always be tightly connected to huge investments in infrastructure. Therefore only few companies will be able to afford the creation of own GenAI applications. However, based on the models auf these few companies, numerous applications and use cases can be established, which also provide a foundation for numerous new companies using the platforms and models of big GenAI players. Chat GPT started providing this by their own app store, where anyone can develop own GPT’s based on Chat GPT and establish business cases.

11. Do you expect the emergence of GenAI systems to trigger the need to adapt EU legal antitrust concepts?

In general, the EU AI Act provides a good basis for upcoming legal questions. However, the field of GenAI continuously evolves, changes, or adapts. Therefore, also questions about the concrete interpretation and adaptation of the AI Act will arise in many cases. Furthermore, a regulation in terms of the GDPR and audits of GenAI products will be required to ensure, that the rights of the people in the EU are not harmed by current or upcoming GenAI products. Due to the expected fact,

that many smaller companies will be acquired by only few big players, also the antitrust topic becomes relevant. There will be similar discussions as we have with the big players for Operating Systems, Cloud Services, Social Media Networks, Content or Streaming Providers, etc. However, from a legal perspective, the current state of the art should be sufficient to also cover the GenAI area.

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

Also for this question, the answer given for question 11 is valid. However, in terms of the investigation tools and practices, there is the problem, that currently, GenAI is not explainable and there is no proven way to formally validate models in terms of consistency, bias, privacy protection. The area of XAI (Explainable AI), which tries to figure out human understandable and mathematically founded ways of explaining how the source input of GenAI systems has been used to produce a given output, still needs to be further researched and explored. Results of this research have to be implemented into tools, and a standardized way of auditing GenAI systems has to be established.