

Submission by Booking.com to the EC's Call for contributions on Generative AI

The European Commission defines 'Generative AI systems' ('GenAI') as '*AI systems that generate, in response to a user prompt, synthetic audio, image, video or text content, for a wide range of possible uses, and which can be applied to many different tasks in various fields*'.¹ Booking.com understands that 'Foundation Models' ('FM') could fall into this category. The UK's Competition & Markets Authority ('CMA') defines FMs as '*a type of AI technology that are trained on vast amounts of data that can be adapted to a wide range of tasks and operations*'.²

For the objective of defining the relevant market for GenAI, Booking.com submits that it is imperative to distinguish between companies that (1) provide GenAI, such as FMs, and (2) use FMs for their product development or internal productivity applications. Accordingly, Booking.com submits that two separate markets can be defined: (1) (upstream) market for 'FM development and supply' services and (2) (downstream) market for 'FM deployment'. Companies that develop FMs can also deploy them in their own products, hence, can be vertically integrated.

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 a FM, that can effectively compete with currently available FMs, requires significant computing power, training data and expertise at the infrastructure level.

Accordingly, FMs specifically require the following components to be built and trained:

1. Hardware for computing power: Thousands of graphics processing units ('GPUs') or similar (e.g. 'TPUs') to train the models and perform computations, possibly distributed over several data centres

¹ European Commission, Press release on , Commission launches calls for contributions on competition in virtual worlds and generative AI (9 January 2024), available at https://ec.europa.eu/commission/presscorner/detail/en/IP_24_85 (last accessed 4 March 2024).

² CMA, AI Foundation Models: Initial Report (18 September 2023), full report (para. 2.2.), available at https://assets.publishing.service.gov.uk/media/65081d3aa41cc300145612c0/Full_report_.pdf (last accessed 5 March 2024).

2. Data, incl. training data, fine-tuning data and data returned to FM applications at inference. To give an indication, Google's Gemini requires '*trillions of words*' in training data.³
3. Machine Learning ('ML') expertise, including engineers, data and research scientists.

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.

Out of these three factors that contribute to successfully training a FM, sufficient computing power is the most important one. Therefore, Booking.com holds the view that missing computing power is the limiting factor for new market entrants into the FM supply market.

FMs currently on the market (such as GPT, Gemini, Llama) are largely trained on the same data. The data got collected from the web by crawling of new or updated web pages which then got scraped.⁴ The applied parameters used by the FMs vary (i.e. the multipliers/weights and additions/biases). However, NVIDIA makes about 70% of GPU chips used for high performance computing and AI and therefore for FMs.⁵ More than half of the 500k produced GPU chips in Q3 2023 reportedly went to Meta or Microsoft.⁶ Moreover, the CMA recently estimated the costs of training a FM of around \$1.3M (for GPT-3) to \$10M (for PaLM).⁷ Training costs for Meta's Llama 2 are estimated at \$20M, which is a considerably higher investment than the estimated \$2.4M in training costs for the predecessor model.⁸ In addition to these costs, companies would need to invest significant resources, such as energy costs, in the fine-tuning of the FM and FM responses. Lastly, due to the necessary compute power, FMs are typically accessible to FM consumers only via a cloud computing provider (e.g. Amazon's AWS, Microsoft's Azure). Cloud service providers offer the computational resources and services (incl. API solutions, virtual ML

³ Google Ireland Ltd., Gemini FAQs, available at <https://gemini.google.com/faq> (last accessed 6 March 2024).

⁴ CMA, AI Foundation Models: Initial Report (para. 2.10.).

⁵ Dylan Sloan, Nvidia shatters stock market record by adding over \$230 billion in value in one day. Here's why it's dominating the AI chip race, Fortune (22 February 2024) available at <https://fortune.com/2024/02/22/what-does-nvidia-do-chips-ai-jensen-huang/> (last accessed 6 March 2024).

⁶ Anton Shilov, Nvidia sold half a million H100 AI GPUs in Q3 thanks to Meta, Facebook — lead times stretch up to 52 weeks: Report, Tom's HARDWARE (28 November 2023) available at <https://www.tomshardware.com/tech-industry/nvidia-ai-and-hpc-gpu-sales-reportedly-approached-half-a-million-units-in-q3-thanks-to-meta-facebook> (last accessed 6 March).

⁷ CMA, AI Foundation Models: Initial Report (para. 2.18.).

⁸ Jonathan Vanian, Meta's unique approach to developing AI puzzles Wall Street, but techies love it, CNBC (16 October 2023) <https://www.cnbc.com/2023/10/16/metas-open-source-approach-to-ai-puzzles-wall-street-techies-love-it.html> (last accessed 6 March 2024).

development environments or development tools) necessary for selecting, testing, deploying, training and running large-scale ML models. Cloud service providers offer scalable infrastructure solutions that allow organisations to deploy ML models based on FMs.

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?

For Booking.com, the current choice of a suitable FM provider largely depends on three factors:

- The first factor is the ability of Booking.com to access different FMs (directly or via API) for different purposes via one cloud services provider. These cloud service providers offer the necessary flexibility in choice of FMs, the compute power and infrastructure options that Booking.com requires to successfully leverage the efficiencies of FMs.
- The second factor is the technical maturity of the developed FM.
- The third factor represents the contracting terms of FM providers that (1) vary substantially in main areas (incl. in relation to liability and indemnification clauses) and (2) can or cannot be negotiated individually.

Having said that, it seems likely that all the three main cloud services providers (AWS, Azure and Google Cloud Platform) will offer similar FM capabilities in the future. Moreover, already today 3rd party providers offer API endpoints to cloud services providers to fetch embeddings (i.e. the numeric representations of data or files) for them to be used in GenAI applications. These 3rd party embedding stores are therefore facilitating moving data between cloud services providers.⁹

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.

As discussed in our response to Question 1, on the FM infrastructure level, we observe that potential competitors of established FM providers would specifically need to get access to the necessary computing power (via cloud services providers) and high-end chips.

On the FM consumer market, we observe that the majority of current FM providers (1) will not individually negotiate their contracting terms and (2) include terms to the clear disadvantage of FM deployers. For example, such terms often expressly:

⁹ Cf. Coralogix' "Open Search API". Coralogix, Open Search API (11 October 2023), available at <https://coralogix.com/docs/opensearch-api/> (last visited 7 March 2024).

- Disclaim all warranties: FM providers disclaim all warranties as to non-infringement, data losses and product fitness for purposes.
- Reserve rights to use the data provided by FM consumers: FM providers reserve all rights to use data inputted by the FM consumer to (i) develop and improve the applicable service and its underlying technologies and (ii) develop and improve affiliate AI technologies that are not available to the FM consumer.
- Reserve rights, title and interest to the output of the FM: FM providers claim the rights to the output of their FM that is used in the products developed by the FM consumer.
- Include non-compete clauses: FM consumers are not allowed to use the FM provider's services (incl. the output of the FM) to develop or improve (i) similar or competing FM services or (ii) *any* other ML model.

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?

Based on Booking.com's current observations, it seems likely that FM providers will continue monetizing via charging monthly licensing fees that are tied to a token-model for inference (based on actual consumption and other variables).

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.

Open source models effectively democratise the production of AI modelling by allowing researchers and developers across industry to work on the same, or similar, models. Rather than being locked-in to a closed source model, an open source model should allow for more innovation as it evolves with public scrutiny. The expectation is that by using open source AI, companies would meet transparency and fairness requirements more easily, as open source models make clear progress in areas like anti-discrimination and safety.¹⁰ The requirements under the EU AI Act ('EU AIA') are different for open source models (which should be assessed based on their contractual terms vis-a-vis FM consumers), at times allowing for exemptions from documentation requirements, which help equalise the playing field with proprietary systems at least during the development stage.

¹⁰ Daniel Castro, Center for Data Innovation, The EU's AI Act Creates Regulatory Complexity for Open-Source AI (4 March 2024) <https://datainnovation.org/2024/03/the-eus-ai-act-creates-regulatory-complexity-for-open-source-ai> (last accessed 6 March 2024).

Booking.com has the view that open source models represent an alternative to closed source models, if these models get fine-tuned for a specific purpose. Open source models are generally more affordable and energy-efficient and can reach performance levels of closed source models.

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 explained in our response to Question 1, we consider access to data a prerequisite for successfully training and fine-tuning of FMs for specific use cases.

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?

Booking.com observes that there are low technical hurdles to switching FMs that are provided via APIs (including API-integrations offered by Open AI, AWS Bedrock or Google's Gemini). Switching cloud services providers, however, would involve considerable effort in time and resources, as this would mean that all proprietary data that is stored in the cloud would need to be moved to the new cloud provider.