

Report Regarding Generative AI ver. 2.0



Table of Contents

Chapter 1. Introduction.....	3
1. Background.....	3
2. Publication of Report Regarding Generative AI ver. 2.0.....	4
Chapter 2. Structure of the Generative AI Market.....	5
1. Infrastructure layer.....	6
(1) Computing resources.....	6
A. Semiconductor chips.....	6
B. AI cloud services.....	9
C. Summary.....	9
(2) Data.....	10
A. Quantity and quality of data.....	10
B. Types of data.....	10
C. Competitive advantage.....	11
D. Summary.....	11
(3) Expertise (Specialized talent).....	12
A. Competition for human resources, including Big Tech companies, etc.	12
B. Competition for human resources among domestic enterprises.....	12
C. Summary.....	13
2. Model layer.....	13
(1) Trends in generative AI models.....	13
(2) Japanese language proficiency of models.....	15
(3) Trends in other generative AI models.....	15
(4) Model development methods.....	15
(5) Summary.....	16
3. Application layer.....	16
(1) Application layer market trends.....	16
(2) Integration with existing digital services.....	17
(3) AI agents.....	18
(4) Summary.....	19
4. Other generative AI-specific issues and considerations that transcend layers.....	19
(1) Switching and migrating development environments, etc.....	19
(2) Open source/closed source.....	19
(3) Partnerships.....	20
(4) Generative AI models on mobile OS.....	22
[Reference] Services that apply generative AI technology to existing fields of technology.....	23
Chapter 3. Antimonopoly Act and Competition Policy Issues Surrounding Generative AI.....	28
1. Restrictive practices concerning dedicated software on mobile OS.....	28
2. Integrating generative AI into existing digital services.....	32
3. Other points of discussion.....	36
(1) Self-preferencing.....	36
(2) Parallel conduct using generative AI.....	36

(3) Acquisition of highly skilled professionals through partnerships.....	36
Chapter 4. Next steps for the Japan Fair Trade Commission	38
(Appendix) Foreign Competition Authorities' Work in Generative AI Space...	40

Chapter 1. Introduction

1. Background

Against the backdrop of rapid technological advancements and growing public interest in recent years, generative AI has been permeating diverse sectors of society. The use of generative AI has had a major impact on the economy and society by promoting operational efficiency and creating new services and business models. In addition, improvements in the performance of generative AI models¹ have expanded the range of content that can be generated, and have become an important basis for creative activities and information dissemination. In addition, advances are being made in research on robotics and autonomous driving technologies applying generative AI technology, driving transformative changes in industrial structures and lifestyles. In this way, generative AI is driving new innovations across a wide range of fields.

On the other hand, various risks have been pointed out with regard to generative AI, including concerns over the infringement of intellectual property rights including copyright, as well as the risk of disrupting society through the proliferation of false or misleading information. Some of these risks are pointed out as risks from a competition policy perspective. The Japan Fair Trade Commission (JFTC) has been conducting a market study to understand the actual conditions of the generative AI markets from the viewpoint of maintaining a fair and free competition environment in Japan's generative AI markets, ensuring the sustainable development of the technology to generate further innovation, and integrating generative AI in a sound manner for the economy and society at large.

In light of the current state of flux in generative AI markets, the JFTC has proceeded with this market study in a more rapid and flexible manner than previous studies. Specifically, in October 2024, the JFTC published a discussion paper titled “Generative AI and Competition” (hereinafter referred to the “discussion paper”)² to solicit information and opinions widely from parties concerned. This paper outlines the market structure of the generative AI market and presents key issues from the perspectives of the Antimonopoly Act and competition policy. Concurrently, the JFTC conducted hearings with various stakeholders including domestic and foreign enterprises. The JFTC swiftly analyzed the trends in the information and opinions collected through the request for comments, hearings, etc., carefully selected major themes, and, in June 2025, compiled the “Report Regarding Generative AI ver.1.0” (hereinafter referred to as the “last report”) in a more targeted form than the last report.³

¹ A general term for the models used in generative AI products (see Chapter 2 below). This includes foundation models and models that have been fine-tuned from them.

² Japan Fair Trade Commission “Discussion paper ‘Generative AI and Competition’” October 2, 2024 (<https://www.jftc.go.jp/en/pressreleases/yearly-2024/October/1002.html>)

³ Japan Fair Trade Commission “About the Report Regarding Generative AI ver. 1.0” June 6, 2025 (<https://www.jftc.go.jp/en/pressreleases/yearly-2025/June/250606.html>)

2. Publication of Report Regarding Generative AI ver. 2.0

As described in 1 above, this survey adopts a faster and more flexible method than conventional market studies. For this reason, the JFTC has continued its market study on the actual conditions of the generative AI market even after the publication of the last report. The JFTC has collected a wide range of information and opinions, including the most recent developments, by establishing a new information submission form concerning generative AI on its website and conducting hearings with approximately 30 stakeholders, including domestic and foreign enterprises, experts, relevant ministries and agencies, and overseas authorities.

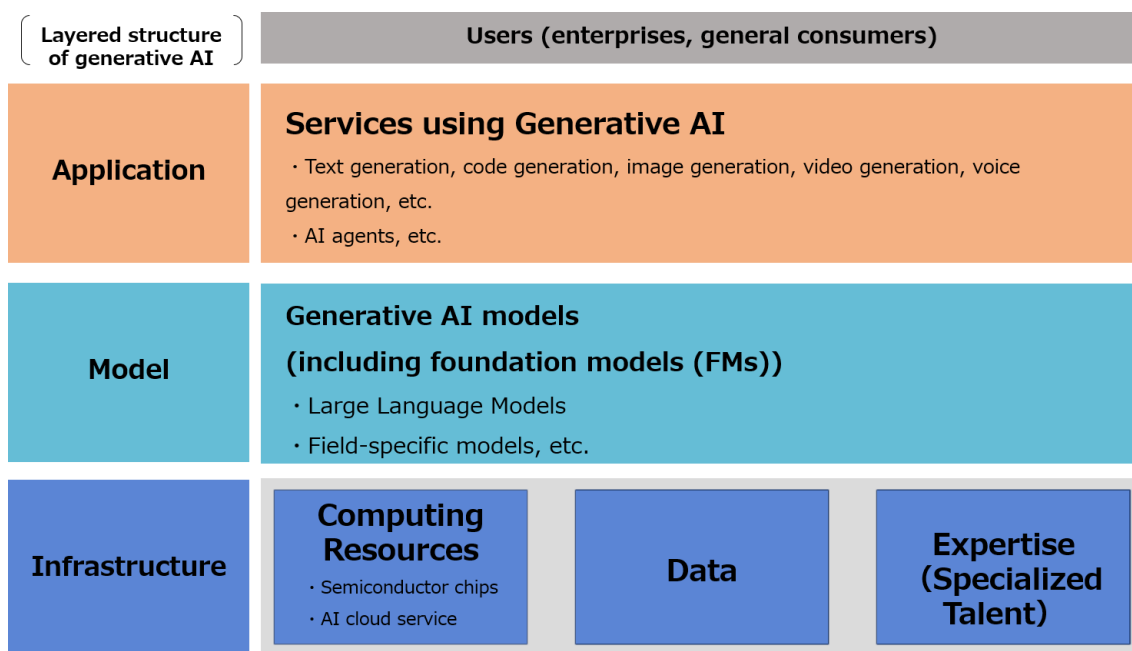
Based on the information, opinions and other materials collected, the JFTC updated (i) the overview of the generative AI market, added (ii) an overview of the autonomous driving market, and reorganized (iii) issues under the Antimonopoly Act, thereby compiling the “Report Regarding Generative AI ver. 2.0” (hereinafter referred as the “report”). The report consolidates the contents of the discussion paper and the last report and re-examines the current actual conditions of the generative AI market, including points that were not fully captured in the last report.

Chapter 2. Structure of the Generative AI Market

Generative AI is an artificial intelligence technology⁴ capable of generating various types of content—such as text, videos, still images, illustrations, and audio—in response to prompts (questions or instructions) and is developed using models that have been fine-tuned⁵ based on pre-trained⁶ foundation models⁷. In addition, applications and services (hereinafter referred to as “generative AI products”) equipped with functions to generate text, videos, still images, illustrations, audio, and other content using this model are provided.

The market structure of the generative AI sector consists of multiple layers, with domestic and international companies conducting business activities at each level. In this report, the generative AI market is organized into three layers: “Infrastructure,” “Models,” and “Applications,” as shown in Figure/ Table 1.

Figure/Table 1. Structure of the Generative AI Market



*AI cloud services will be discussed later.
Source: Prepared by the JFTC.

⁴ Act on Promotion of Research and Development, and Utilization of AI-related Technology (Act No. 53 of 2025) Article 2. The term “artificial intelligence-related technology” as used in this Act means technology necessary for realizing the function to substitute intellectual abilities for human cognition, inference and judgement by artificial means, and technology relating to information processing systems for realizing the function to process the input information and output that result by using the technology.
⁵ Additional training performed on a pre-trained model to optimize it for a specific task or field. The model’s performance is further improved by tailoring it to specific business domains or applications.
⁶ The process of initial training with large amounts of data during the development of the foundation model. For example, pre-training on text data may involve the use of billions to tens of trillions of tokens (words and phrases). This process creates a foundation on which the model can capture a broad range of knowledge and apply it to diverse tasks.
⁷ Foundation Model (FM) is the core technology that supports generative AI and is the basis for creating individual models. The foundation model is created by pre-training using a large amount of data. The foundation model can be optimized for a specific task or field by fine tuning.

1. Infrastructure layer

The infrastructure layer is the market that supports generative AI. The three main elements of this layer are computing resources, data, and expertise (specialized talent).

(1) Computing resources

The development of generative AI requires semiconductor chips composed of dedicated circuits for high-speed parallel processing, such as GPUs⁸, which are hereinafter simply referred to as “semiconductor chips”⁹. There are two methods for utilizing semiconductor chips as computing resources: (1) procuring them for use directly from semiconductor chip providers, and (2) using services that provide the performance of semiconductor chips via the cloud (hereinafter referred to as “AI cloud services”).

The following section provides an overview of the markets for semiconductor chips and AI cloud services.

A. Semiconductor chips

The market situation for semiconductor chips in the generative AI sector differs depending on whether they are used in the training phase or the inference phase of generative AI models.

(a) Training phase

The training phase refers to the stage in the process where a generative AI model uses large amounts of data to identify common rules and features, enabling it to generate appropriate responses and content based on those findings.

There is a tendency in the performance of generative AI models known as the “scaling laws,” whereby performance improves as the number of model parameters¹⁰, the volume of data, and the computational load increase. For this reason, a sufficient supply of semiconductor chips during the training phase is considered essential for developing high-performance generative AI models.

While there are many types of semiconductor chips, GPUs, which were originally designed for image processing, have high parallel processing capabilities and are considered ideal for developing

⁸ A GPU (Graphics Processing Unit) is a type of semiconductor chip specialized for screen display and image processing. Because it excels at parallel processing—the ability to process large amounts of data simultaneously—it is considered well-suited for tasks such as the development of generative AI.

⁹ This semiconductor chip is sometimes collectively referred to as an AI accelerator.

¹⁰ The number of variables that make up a generative AI model. The foundation model is used to process input data. While it is generally believed that an increase in the number of parameters improves the performance of generative AI models, there are also drawbacks, such as increased training time and costs, as well as higher computational load during inference.

generative AI models. In this GPU market, NVIDIA GPUs account for the majority of the global market share. The reasons why NVIDIA GPUs hold such a large market share include: (1) high performance (fast computation), (2) a well-established development environment, and (3) the company's ongoing commitment to allocating resources to improve technologies such as those mentioned in points (1) and (2). In particular, NVIDIA's development environment (such as CUDA¹¹) is said to offer high usability and scalability because deep learning frameworks like PyTorch and TensorFlow¹² use CUDA libraries¹³ to be highly optimized for NVIDIA GPUs. In addition, these NVIDIA development environments are generally compatible only with NVIDIA GPUs.

In the hearings conducted with business operators and others in preparation for the publication of this report (hereinafter referred to as "the current hearings"), many respondents noted that NVIDIA GPUs continue to dominate the market, particularly in the training phase. Additionally, regarding the supply and demand situation for NVIDIA GPUs, while there was a period when supply improved, the situation has recently become tight again. These observations suggest that demand for NVIDIA GPUs in the training phase remains high.

(b) Inference phase

The inference phase refers to the stage in the process in which a model that has undergone the training phase analyzes new input data and generates optimal responses or content based on the patterns and knowledge acquired during training.

During the inference phase, there is a need for semiconductor chips that can process data quickly and with low power consumption in order to perform efficient inference; generally speaking, the computational capacity required during this phase is not as great as that needed during the training phase. As for the semiconductor chips used in the inference phase, in addition to NVIDIA GPUs, chips developed by Big Tech companies, established semiconductor manufacturers, and startups are also attracting attention in the market (see (c) below). Furthermore, driven by the emergence of technologies such as Test-Time Scaling¹⁴ and models capable of

¹¹ CUDA is a parallel computing platform and programming model developed by NVIDIA, enabling general-purpose computing using GPUs.

¹² A deep learning framework refers to software designed to efficiently build, train, and perform inference on neural networks.

¹³ Software libraries used on CUDA such as cuDNN. Note that a software library refers to a collection of files that compiles functions and program components frequently used in programming into a format that is easy to reuse. <https://biz.kddi.com/content/glossary/l/library/>

¹⁴ Unlike traditional methods that generate responses to prompts immediately, this is a general term for technologies that improve the accuracy of generative AI model responses by allocating more computational resources to inference.

handling long contexts¹⁵, as well as the growing number of users of generative AI services, demand for semiconductor chips used in the inference phase is also increasing. While competition in the inference chip market is intensifying, NVIDIA GPUs—which continue to hold an advantage in development environments and other areas—remain in high demand even at the inference stage, at the current moment.

During the current hearings, opinions were expressed that while new semiconductor chips for inference are being developed one after another and competition in the inference phase is intense, NVIDIA GPUs remain dominant. Other opinions cited reasons for NVIDIA GPUs' continued dominance in the inference phase, such as the fact that software environments are often designed with NVIDIA GPUs in mind, leading to additional costs when adopting semiconductor chips other than NVIDIA GPUs.

(c) Semiconductor chips other than NVIDIA GPUs

As noted in (a) and (b) above, while there is strong demand for NVIDIA GPUs in both the training and inference phases, semiconductor chips other than NVIDIA GPUs have also been gaining prominence in recent years. Starting with Google's TPU (Tensor Processing Unit) and Amazon's Trainium, Microsoft and Meta are also actively developing their own semiconductor chips in an effort to optimize their infrastructure and reduce their reliance on NVIDIA GPUs. In addition to improvements in computing power, these semiconductor chips are also seeing software enhancements, such as support for deep learning frameworks. In addition, OpenAI, for example, announced that it had entered into a strategic partnership with AMD in October 2025 and would receive 6 gigawatts of AMD GPUs from the company starting in 2026¹⁶. Furthermore, OpenAI announced that it had entered into a long-term agreement with Broadcom in October 2025 regarding the joint development and supply of semiconductor chips, and that it plans to deploy server racks incorporating 10 gigawatts worth of semiconductor chips—jointly developed by the two companies—at OpenAI facilities and the data centers of OpenAI's partner companies from the second half of 2026 through 2029¹⁷. In this way, supply of semiconductor chips other than NVIDIA GPUs is also expected to increase in the future, and the market share of such chips may grow in the years to come.

¹⁵ Generative AI model with large context window (context length). The context window (context length) is the amount of text per token that the generative AI model can store at one time. Enlarging the context window allows the model to process longer inputs, thereby improving the accuracy of the generative AI model's responses. <https://www.ibm.com/jp-ja/think/topics/context-window>

¹⁶ On October 6, 2025, AMD and OpenAI announced that they had entered into a strategic partnership regarding the rollout of AMD GPUs. <https://openai.com/index/openai-amd-strategic-partnership/>

¹⁷ On October 13, 2025, OpenAI and Broadcom announced a strategic partnership and revealed plans to roll out 10-gigawatt semiconductor chips designed by OpenAI. <https://openai.com/ja-JP/index/openai-and-broadcom-announce-strategic-collaboration/>

On the other hand, during the current hearings, some respondents noted that while technological innovation in semiconductor chips is rapid, once a new chip is introduced, switching to a different one incurs significant costs and, as a result, many AI cloud service providers are currently weighing the optimal timing for introducing new semiconductor chips. Furthermore, many believe that NVIDIA GPUs remain the dominant choice at this point, and it appears that NVIDIA's dominance will continue for the foreseeable future.

B. AI cloud services

Access to proprietary large-scale computing resources is primarily limited to major companies such as Big Tech firms. As a result, the majority of developers of generative AI models and products use AI cloud services. By offering these services, developers of generative AI models and products are now able to train and run inference on computationally intensive generative AI models. Against this backdrop, as demand for generative AI grows, so does the demand for GPU servers, and competition is intensifying in the AI cloud services market as various cloud service providers seek to differentiate themselves.

Meanwhile, the three major cloud service providers—Amazon Web Services (AWS), Microsoft Azure, and Google Cloud—are leveraging their substantial financial resources to secure semiconductor chips and offer AI cloud services optimized for generative AI development; as a result, these three companies continue to hold a significant share of the Japanese market for AI cloud services. Furthermore, it has been pointed out that the barriers to entry are high because significant investment is required in the data centers that operate these systems and in the cooling systems needed to handle the enormous amounts of heat they generate¹⁸. Therefore, these three major cloud service providers are expected to continue to dominate the market.

C. Summary

In the GPU market, NVIDIA GPUs hold a significant share of the global market due to factors such as development environments like CUDA. Particularly in the training phase, as mentioned in point A. above, a sufficient supply of semiconductor chips is essential, and NVIDIA GPUs are said to continue to hold a competitive advantage. In contrast, in the inference phase market, competition has intensified compared to the training phase—driven by factors such as the emergence of new entrants—and innovation is advancing through the efforts of a diverse range of players.

¹⁸ FTC "Partnerships Between Cloud Service Providers and AI Developers" January 17, 2025, p. 10 (<https://www.ftc.gov/news-events/news/press-releases/2025/01/ftc-issues-staff-report-ai-partnerships-investments-study>)

Furthermore, while semiconductor chips other than NVIDIA GPUs are gaining ground and efforts to reduce dependence on NVIDIA are gaining momentum, NVIDIA is expected to remain the dominant player in both the training and inference phases for the foreseeable future.

In the AI cloud services market, the three major cloud service providers are expected to continue to hold strong positions.

(2) Data

A. Quantity and quality of data

The quantity and quality of the data used vary depending on the phases of generative AI model development—such as pre-training, fine-tuning, and RLHF¹⁹—as well as subsequent use scenarios like RAG²⁰, as described below.

Pre-training requires a large amount of data in order to learn language structures and foundational knowledge. In fine-tuning and RLHF, data tailored to specific objectives—such as improving the accuracy of responses to user instructions and questions—is required. For data used in RAG, it must be organized in such a way that the necessary information can be retrieved appropriately. Furthermore, when developing specialized models tailored to specific industries or applications, such as finance or healthcare (see Section 2(1) below), high-quality data specific to such domains is required.

B. Types of data

There are various types of data available for developing generative AI models, including data published on the web, publicly available datasets, data held by private companies, and data held by government agencies. However, while the development of general-purpose large language models²¹ requires vast amounts of data, much of the high-quality data accessible on the internet is already being used to train generative AI, leading to concerns that open data available for training may eventually run out²². In this context, there has been a growing trend toward using synthetic data²³ to compensate for a lack of training data, as well as toward improving the quality of training data to make the most of limited

¹⁹ RLHF (Reinforcement Learning from Human Feedback) is reinforcement learning based on human evaluation.

²⁰ RAG (Retrieval-Augmented Generation) is defined as a language generation model that combines pre-trained parametric and non-parametric memory (i.e., retrieval-based memory) (Lewis, Patrick et al. “Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks” *Advances in Neural Information Processing Systems* 33 [2020]: 9459-9474). In corporate settings, it is used not only to search internal documents and databases to enhance the accuracy of AI-generated responses, but also to search for information on the internet in real time so that the most up-to-date data can be provided in responses.

²¹ Large Language Model (LLM) is a language-specific generative AI model. It has been pre-trained on a massive amount of text data and is capable of generating text—such as conversations and academic papers—just like a human.

²² In an article published on December 11, 2024, the academic journal “Nature” summarized the current state of AI development and pointed out that training data could run out by 2028 (Nature “The AI revolution is running out of data. What can researchers do?” December 11, 2024 (<https://www.nature.com/articles/d41586-024-03990-2>)).

²³ Synthetic data refers to artificial data generated by computer algorithms that closely resemble real-world data.

resources.

During the current hearings, participants noted that the burden of individually verifying licensing status when collecting data is significant, leading many to utilize synthetic data while purchasing external data as needed. Others pointed out that relying solely on synthetic data for repeated training carries the risk of compromising model performance, and emphasized that it will be crucial in the future to utilize the vast amounts of internal corporate data that have been accumulated but remain underutilized.

C. Competitive advantage

Some observers point out that, when it comes to data collection, Big Tech companies have a competitive advantage because, in addition to being able to purchase vast amounts of data thanks to their substantial financial resources, they can also obtain data through the services they provide.

During the current hearings as well, opinions were expressed that Big Tech companies are able to obtain vast amounts of various data from their own services, and that these data are contributing to the improved performance of the generative AI models they develop.

On the other hand, some overseas model developers expressed the view that, overall, only a very small fraction of the data collected through their services is actually used to train their own generative AI models. Furthermore, given that many companies developing generative AI models do so without utilizing data obtained through services other than their own, it was argued that such data is not necessarily indispensable for the development of generative AI models. In addition, many domestic companies also noted that, even if they were able to obtain the same volume of data as Big Tech firms, they do not possess the computing resources (such as data centers or supercomputers) required to process it. Consequently, they do not necessarily require the same volume of data and are compelled to adopt different strategies.

D. Summary

The demand for data varies depending on the context and purpose, and while quantity is important, quality is sometimes prioritized as well. Against this backdrop, concerns have been raised that the data available for training may eventually run out. As a result, efforts are underway to utilize synthetic data, and there is growing interest in leveraging internal corporate data that has been accumulated but not fully utilized. Furthermore, while some argue that Big Tech companies hold a competitive advantage in data acquisition, others contend that this advantage is limited. We believe that assessments of competitive advantage will vary depending on the types of models that generative

AI model developers aim to create and the computational resources they possess²⁴.

(3) Expertise (Specialized talent)

A. Competition for human resources, including Big Tech companies, etc.

The rapid development and market expansion of generative AI technology has led to a significant increase in the demand for highly skilled professionals in the development of generative AI, and the highly specialized researchers, engineers, data scientists, and governance professionals who support the development and application of generative AI models and products are an essential driver of industry growth. In particular, Big Tech companies are noted to have an advantage in the competition for highly skilled professionals backed by their financial strength, computational resources, and other assets.

At the current hearings, there was an opinion that it is extremely difficult for domestic operators to acquire specialized personnel on the same terms as Big Tech companies because Big Tech companies can offer high remuneration levels and have overwhelmingly large computational resources.

However, it has been noted that Big Tech companies do not always have a monopoly on highly skilled professionals, and that there is a certain degree of mobility of such human resources. For example, there are examples of engineers who start their own businesses or leave Big Tech companies to take up new positions at startups or other companies. During the current hearings, there were comments noting an increasing trend of talented individuals seeking to make their mark moving from Big Tech companies to startups or starting their own businesses.

B. Competition for human resources among domestic enterprises

Such mobility of highly skilled professionals is thought to occur mainly in Big Tech companies, and not all companies, including domestic enterprises, are benefiting from it in the same way. Challenges faced by domestic enterprises in securing highly skilled professionals include lower compensation levels and lack of computational resources compared to Big Tech companies. On the other hand, domestic companies are believed to have the advantage of being able to more easily train and hire local human resources.

During the current hearings, while some participants noted that competition for human resources is becoming increasingly fierce as overseas generative AI startups expand into Japan—despite the limited

²⁴ Since the topic of “Processing of Rights, etc. in The Collection and Use of Data,” which was addressed in the last report, is partially related to the follow-up survey on the fact-finding survey on the news content distribution sector launched by the JFTC in December 2025, we plan to examine this matter within the scope relevant to that follow-up survey.

pool of highly skilled professionals in the country—others pointed out that employees at Big Tech companies tend to have relatively short tenures, and that human resources seeking long-term employment with a single company are more likely to settle with domestic firms.

C. Summary

As the demand for highly skilled professionals in generative AI continues to rise, competition for such human resources is becoming more intense, both domestically and internationally. In this environment, while Big Tech companies hold an advantage in the competition for highly skilled professionals due to their financial strength and abundant development resources, the competition for human resources—both among Big Tech companies themselves and between Big Tech companies and startups—remains intense. On the other hand, while domestic companies face challenges such as low pay levels and a shortage of computing resources, they also have unique strengths, such as the ability to easily train and hire local human resources.

2. Model layer

The model layer represents a market focused on developing and providing generative AI models, which are pre-trained using vast amounts of data.

(1) Trends in generative AI models

The main applications of generative AI models can be classified into three main categories: (1) text generation, (2) image (video and 3D model) generation, and (3) voice generation. Generative AI models can be classified into general-purpose models that can be applied to a variety of tasks and specialized models for specific applications, such as finance and healthcare.

Amid this situation, the race to develop large language models continues to intensify, particularly among Big Tech companies and others²⁵. These large language models are essentially models that each company has developed in-house from scratch, with a primary focus on versatility. While some domestic companies are developing general-purpose large language models, the development process generally requires significant investment in computing resources, the collection of vast amounts of data, a large number of highly skilled professionals, and a long development cycle; as a result, Big Tech companies with substantial financial and technical resources are said to have an advantage.

²⁵ In February 2026, for example, Anthropic launched the “Claude Opus 4.6” and “Sonnet 4.6,” and Google launched the “Gemini 3.1 Pro.” In addition, each company is continuously updating and introducing new models, with Google releasing the “Gemini 3.1 Flash-Lite,” and OpenAI launching the “GPT -5.4,” “GPT -5.4 mini,” and “GPT -5.4 nano” in March of the same year.

During the current hearings, it was noted that while the model performance of Big Tech companies continues to improve and that of Chinese providers is also steadily advancing—indicating that competition in model development remains vigorous—the performance gap between general-purpose models is narrowing. Consequently, the competitive landscape is converging to some extent, and the industry is shifting from a phase where new models are released one after another to one where multiple high-performance models are continuously refined.

On the other hand, there are examples—primarily among domestic companies and startups—of firms seeking to differentiate their products from Big Tech companies’ general-purpose models by utilizing generative AI models provided by other companies (such as open-source models²⁶) and adapting them for specific applications (such as healthcare, finance, law, and manufacturing). However, it is said that Big Tech companies, which possess high-performance foundational models, still hold an advantage even in the development of such specialized models.

During the current hearings, opinions were expressed that while domestic companies’ competitive edge lies in developing domain-specific models that leverage proprietary data—such as data not available on the web—there is also a possibility that Big Tech companies’ general-purpose models could outperform them across a wide range of fields, including specialized domains. Additionally, it was noted that while high Japanese language performance had once been considered a strength of domestic companies, there will be an increasing demand in the future for the accuracy of domain-specific specialized models and the high level of trust they inspire among customers.

Against this backdrop, as noted in the opinions mentioned above, the development of Chinese-made generative AI models is also gaining momentum. In January 2025, Chinese startup DeepSeek released a large language model (DeepSeek-R1). Although safety concerns had been reported regarding this model, it is said to offer performance on par with the latest models from other companies, and this had a significant impact on the U.S. stock market and other markets²⁷. Since then, Chinese models other than those from DeepSeek have also emerged, and some domestic companies are now using these models.

On the other hand, there are still concerns about the use of these models. During the current hearings as well, opinions were expressed that some domestic operators remain cautious about using Chinese-made models from the perspectives of economic security and ensuring the safety of outputs.

²⁶ “Open source” refers to software the technical specifications of which are made publicly available, allowing anyone to freely modify and optimize it. In contrast, “closed source” refers to software the technical specifications of which are not made publicly available and are accessible only to specific individuals.

²⁷ DeepSeek has also been cited as one of the factors influencing companies’ demand forecasts (Nihon Keizai Shimbun “A turning point for NVIDIA’s ‘dominant position’” February 27, 2025 (<https://www.nikkei.com/article/DGXZQOGN260GV0W5A220C2000000/>)).

(2) Japanese language proficiency of models

One of the important factors in generative AI models used in Japan is Japanese language proficiency. In the past, there have been reports of foreign-developed models producing unnatural Japanese, and as a result, domestic companies have increasingly prioritized the development of models that emphasize Japanese language proficiency.

During the current hearings, it was noted that while the gap in Japanese language proficiency between overseas and domestic models is gradually narrowing in general-purpose domains, there remains room for domestic companies to maintain a competitive edge in specific specialized fields or areas requiring advanced Japanese language skills by training their models using proprietary Japanese data.

(3) Trends in other generative AI models

In addition to large language models used for text generation, competition is intensifying in the development of generative AI models capable of generating images (including videos and 3D models) and audio, with new models being announced. Models other than large language models generally face challenges similar to those encountered in the development of large language models.

Furthermore, as the scope of generative AI applications has expanded significantly in recent years, multimodal AI—which integrates and processes different types of data such as text, images, audio, and video—has become firmly established. Multimodal AI is expected to expand the scope of AI applications in fields such as healthcare, drug discovery, education, and entertainment, as well as improve processing capabilities for general tasks such as inference and analysis²⁸.

Such technological advancements will further expand the scope of generative AI applications, intensifying competition in the fields of multimodal AI and related products, and are likely to continue contributing to the growth of the generative AI market.

(4) Model development methods

Behind the highly accurate technology of generative AI lies the challenge of requiring huge models, large numbers of parameters, and vast amounts of computational resources. In addition, in order to enable operation in resource-constrained environments such as on mobile devices, there is a need to make models smaller and inference more efficient. To address

²⁸ The Ministry of Internal Affairs and Communications and the Ministry of Economy, Trade and Industry “AI Guidelines for Business Ver1.2 Appendix” March 31, 2026, p. 17 (https://www.soumu.go.jp/main_content/001064286.pdf
https://www.meti.go.jp/shingikai/mono_info_service/ai_shakai_jisso/pdf/20260331_3.pdf)

these challenges, approaches such as “distillation”²⁹ and “Mixture of Experts (MoE)”³⁰ have been proposed³¹ with the aim of making models smaller, lighter, and more efficient, although these methods are not necessarily new.

(5) Summary

In the model layer field, there is intense competition to develop general-purpose models, particularly among Big Tech companies. In response, domestic companies are seeking to differentiate themselves from Big Tech companies’ general-purpose models by developing models tailored to specific applications. In particular, while Japanese language proficiency is crucial for models used in Japan, as the gap in Japanese proficiency between foreign and domestic models narrows, it has been pointed out that domestic operators may gain a competitive edge in specialized fields or situations requiring advanced Japanese language skills.

3. Application layer

The application layer represents a market where generative AI products are developed and offered. Generative AI products for various applications such as text, code, images, video, voice and music are being developed and offered by many companies including startups.

(1) Application layer market trends

Generative AI models used by generative AI product developers can be divided into 1) those provided as open source, 2) those provided as closed source, and 3) those developed for in-house use. Generative AI products are widely deployed across various industries and sectors, including finance, construction, healthcare, and law, and are often developed and delivered through cloud services provided primarily by Big Tech companies.

In the application layer, a wide range of companies—from Big Tech companies to startups—are offering various generative AI products, and there is fierce competition to meet the diverse range of demand. In Japan as well, various generative AI application services—including chatbots for

²⁹ Distillation (knowledge distillation) is generally a method of transferring the wealth of knowledge learned by a pre-trained generative AI model (teacher model) to another generative AI model (student model) that is smaller and more efficient. Typically, a teacher model requires a huge amount of parameters and computational resources, while a distilled student model is said to have a major advantage in that it can achieve the same or similar performance as a teacher model in a low-cost, fast inference environment. Therefore, distilled student models are widely used as optimization methods, especially for edge AI for smartphones.

³⁰ Mixture of Experts (MoE) is a method of demerging a single huge model into multiple subnetworks of “experts” and generating output by selecting the most appropriate expert for each input. This method reduces the computational costs associated with training and inference of the generative AI model, and achieves larger scale, lower power consumption, and higher speed. Through this mechanism, each expert is said to be able to specialize in a particular area or feature of the input, and the amount of computation is said to be reduced while maintaining high expressive power as a whole.

³¹ For example, DeepSeek—a company said to have developed a low-cost, high-performance generative AI model—has announced that it is incorporating distillation and MoE techniques to optimize computational efficiency (https://github.com/deepseek-ai/DeepSeek-R1/blob/main/DeepSeek_R1.pdf). These efficiency-enhancing measures are believed to have the potential to reduce reliance on semiconductor chips such as GPUs to a relative degree.

internal and external use, search systems, text generation, image generation, and speech generation—have been commercialized, and their adoption in the market is gaining momentum. For example, according to OpenAI, the number of ChatGPT users in Japan is reported to have doubled year-on-year in 2024, reaching 6 million³². On the other hand, data³³ suggests that the adoption of generative AI by domestic companies remains low compared to other major countries.

During the current hearings, many respondents noted that the psychological barriers for domestic companies to adopt generative AI products have been lowering. One factor cited as the reason for this is that, as generative AI has become part of everyday life, businesses have gained a deeper understanding of it, and resistance to its introduction is gradually diminishing. That said, while the barriers to adopting generative AI for domestic businesses have lowered, some felt that they still lag behind North America.

(2) Integration with existing digital services

At the application layer, there is a growing trend toward integrating generative AI products with existing digital services—such as search services, office suites, cloud services, social media platforms, and digital advertising tools—and making them available to users. Details are shown in Figure/Table 2.

³² The number of ChatGPT users worldwide reached 400 million in February 2025. As of December 2024, the number of users stood at 300 million. While it took three months for the user base to grow from 200 million to 300 million, the increase from 300 million to 400 million took just over two months, indicating that the pace of user growth is accelerating (Nihon Keizai Shimbun “OpenAI ChatGPT users reach 400 million worldwide, up 30% in 2 months” February 21, 2025 (<https://www.nikkei.com/article/DGXZQOGN20EDF0Q5A220C2000000/>)).

³³ Based on a survey conducted in four countries—Japan, the United States, Germany, and China—we have summarized the current state of AI adoption in the corporate sector. (Omitted) When asked about the use of generative AI in tasks where its application is anticipated, 55.2% of respondents in Japan indicated that they were using generative AI in some form of work (the percentage of respondents who answered “currently using generative AI for business”). With regard to specific tasks, for example, 47.3% of respondents in Japan reported using generative AI to “assist with emails, minutes, document preparation, etc.” (among those who answered “currently using generative AI for business”). In both cases, the ratio remains low when compared with other countries (the Ministry of Internal Affairs and Communications “Information and Communications in Japan White Paper 2025” July 2025 (<https://www.soumu.go.jp/johotsusintokei/whitepaper/ja/r07/html/nd112220.html#f00045>)).

Figure/Table 2. Examples of Integration of Existing Digital Services and Generative AI

Company name	Existing digital services	Integration overview
Google	Google Search	Integration of “AI Overviews” and “AI Mode” with Google Search.
	Google Workspace	Addition of Gemini to Google Workspace (office suite).
	Android OS	Installation of “Gemini” on Android OS devices.
	Chrome	Addition of “Gemini” to the Chrome browser.
Microsoft	Bing Search	Addition of “Copilot” to Bing Search.
	Microsoft 365	Addition of “Copilot” to Microsoft 365 (office suite).
	Windows OS	Addition of “Copilot” to Windows OS devices.
	Edge	Addition of “Copilot” to the Edge browser.
Apple	iOS, etc.	Addition of Apple Intelligence to in-house OS (iOS, etc.) devices.
X	X (SNS)	Addition of X (SNS) to Grok.

Source: Prepared by the JFTC based on various publicly available materials.

Integrating existing digital services with generative AI products has the potential to enhance user convenience by improving the performance of existing digital services. On the other hand, Big Tech companies have established strong positions in specific markets and sectors through the digital services they provide, and integrating generative AI products with their existing digital services has the potential to enhance the competitiveness of those services and further strengthen their market positions (for issues related to the Antimonopoly Act and competition policy, see Section 3.2 below).

(3) AI agents

While 2025 has been dubbed the “year one of AI agents,” and these agents³⁴ have garnered attention as a new technology, their adoption has since gained momentum and they are becoming firmly established in everyday tasks and services. The adoption of AI agent services is gaining momentum in Japan as well³⁵, with initiatives aimed at improving operational efficiency and creating new business models becoming increasingly widespread.

³⁴ In this report, the term “AI agent” refers to an AI system that perceives its environment and makes autonomous decisions in order to achieve specific objectives.

³⁵ For example, SoftBank has announced that it will integrate OpenAI’s AI product “Frontier” into the platform it is currently developing as part of its efforts to commercialize AI agents, and the two companies are collaborating on the development of AI agents under the name “Crystal Intelligence.” (Nihon Keizai Shimbun “SoftBank taps OpenAI for business AI” February 6, 2026 (<https://www.nikkei.com/article/DGXZQOUC063QE0W6A200C200000/>)).

During the current hearings, opinions were expressed that AI agents are expected to advance rapidly in areas such as online shopping and software development coding, and that competition in the AI agent sector—including platforms for AI agents—is intensifying.

(4) Summary

The application layer is one of the most competitive segments within the generative AI market. A wide range of players, from Big Tech companies to startups, have entered the market, and the use of generative AI products in Japan has become more widespread than ever before. In addition, the integration of generative AI products with existing digital services is progressing, and AI agents are also becoming increasingly common in day-to-day operations and services.

4. Other generative AI-specific issues and considerations that transcend layers

(1) Switching and migrating development environments, etc.

In the development of generative AI models, while it is technically possible to switch or migrate development environments—such as by changing hardware or cloud infrastructure—this entails costs and man-hours associated with rebuilding the entire system, re-installing software, and migrating data. For example, switching from one cloud service to another in the development of generative AI models and products will entail costs associated with data migration and other factors. Furthermore, if a particular semiconductor chip is available only through a specific cloud service, it is said that comprehensive switching costs and other expenses will increase because it also requires changing or switching the specific semiconductor chip in use.

Thus, while switching development environments does entail some switching costs and other expenses, whether these constitute a substantial barrier to switching depends on each provider's specific development environment and other factors. During the current hearings, various opinions were expressed, including that many domestic developers tend to rely heavily on existing tools provided by major cloud service providers, resulting in high costs for switching cloud environments; that while the cost of switching development environments for generative AI is not significant, a notable price difference is necessary for switching; that while switching cloud environments is not technically difficult, outsourcing is necessary if a company cannot build its own GPU clusters³⁶ in-house, which consequently increases the switching costs.

(2) Open source/closed source

Generative AI models are available in both open-source and closed-

³⁶ A system in which multiple GPU servers are connected via a network and the GPUs installed in each server are operated in coordination. <https://www.nttpc.co.jp/gpu/gpu-cluster/basic-knowledge.html>

source versions. Open source is expected to lower the barriers to entry for the development and use of generative AI products, particularly for new entrants and startups that struggle to bear high costs, thereby helping to foster a competitive environment. It is also said that because anyone can review and improve the technical specifications, this makes it easier to verify security issues and helps accelerate innovation. However, since the technical specifications have been made public, it has been pointed out that it is difficult for operators to maintain a competitive advantage, and there is also a risk that the technology could easily be exploited for malicious purposes.

On the other hand, one characteristic of closed-source systems is that technical specifications are not disclosed, allowing businesses and research institutions to independently manage and restrict the use of generative AI, thereby reducing the risk of misuse by third parties. However, concerns have been raised about the difficulty of accepting feedback and contributions from third parties, the potential for innovation to stagnate, and the lack of technology transparency due to the lack of disclosure of technical specifications.

Given the above circumstances, while it is difficult to say definitively whether open-source or closed-source approach is preferable from a competition policy perspective, it is considered important to ensure a diverse range of options in the development and provision of generative AI models and products.

(3) Partnerships

In the generative AI market, partnerships are being actively formed both domestically and internationally, including between semiconductor chip providers and generative AI model developers, among generative AI model developers themselves, and between generative AI model developers and generative AI product developers. As Figure/Table 3 shows, there have been several examples of partnerships between Big Tech companies providing existing digital services and startups developing generative AI models.

Figure/Table 3. List of Partnerships

Companies related to generative AI services

	OpenAI	Anthropic	Hugging Face	SakanaAI	
Investor companies	NVIDIA	Alliance ³⁷	Alliance ³⁸ ○ August 2023	○ September 2024	
	AMD	Alliance ³⁹	Alliance ⁴⁰		
	Microsoft	○ July 2019 ⁴¹	Alliance	Alliance ⁴²	
	Amazon	Alliance ⁴³	○ September 2023	○ August 2023	
	Google		○ February 2023	○ August 2023	Alliance ⁴⁴
	Softbank Group	○ February 2025 ⁴⁵			

*○ indicates a partnership involving capital contributions, etc.

Source: Compiled by the JFTC based on various publicized information as of the end of March 2026.

³⁷ On January 21, 2025, the company announced the Stargate Project and revealed that NVIDIA would be a key early technology partner. <https://openai.com/ja-JP/index/announcing-the-stargate-project/>

Furthermore, on September 22, 2025, NVIDIA and OpenAI announced that they had entered into a strategic partnership to build and deploy AI data centers. <https://nvidianews.nvidia.com/news/openai-and-nvidia-announce-strategic-partnership-to-deploy-10gw-of-nvidia-systems>

³⁸ On November 19, 2025, Microsoft, NVIDIA, and Anthropic announced a strategic partnership and revealed that Microsoft and NVIDIA would be investing in Anthropic. <https://news.microsoft.com/source/asia/features/microsoft-nvidia-and-anthropic-announce-strategic-partnerships/?lang=ja>

³⁹ On October 6, 2025, AMD and OpenAI announced that they had entered into a strategic partnership regarding the rollout of AMD GPUs. <https://openai.com/index/openai-amd-strategic-partnership/>

⁴⁰ On June 13, 2023, Hugging Face announced that AMD had joined its hardware partner program. <https://huggingface.co/blog/huggingface-and-amd>

⁴¹ On October 28, 2025, Microsoft and OpenAI announced the expansion of their partnership. Under this partnership, Microsoft announced that it will hold approximately 27% of OpenAI's shares following OpenAI's capital restructuring. <https://openai.com/ja-JP/index/next-chapter-of-microsoft-openai-partnership/>

⁴² On May 24, 2023, Microsoft and Hugging Face announced a partnership to launch the Hugging Face Model Catalog on Azure. <https://huggingface.co/blog/hugging-face-endpoints-on-azure>

⁴³ On November 3, 2025, AWS and OpenAI announced that they had entered into a partnership regarding OpenAI's use of AWS computing resources. <https://openai.com/ja-JP/index/aws-and-openai-partnership/>

⁴⁴ On January 23, 2026, Google and SakanaAI announced that they had entered into a strategic partnership and that SakanaAI would acquire funding from Google. <https://sakana.ai/google/>

⁴⁵ On December 31, 2025, SoftBank Group Corp. announced the completion of an additional investment in OpenAI. <https://group.softbank/news/press/20251231>

In such partnerships, startups developing generative AI models can secure development funding, access the resources needed for development, and access existing related technologies through large-scale investment. Furthermore, for Big Tech companies, licensing innovative generative AI models from startups through partnerships has the potential to enhance user convenience.

On the other hand, it has been pointed out that forming such partnerships could strengthen the position of Big Tech companies and others in the generative AI market, potentially leading to weakened competition⁴⁶ ⁴⁷.

(4) Generative AI models on mobile OS

In recent years, smartphones have become a daily necessity for consumers and serve as the primary point of access to a wide variety of digital content and services. With regard to generative AI technology, advancements such as the miniaturization of generative AI models have made it possible to run these models on some smartphones. These generative AI models that run on smartphones (hereinafter referred to as “on-device generative AI models”) complete all processing within the device via the mobile OS. As the development of on-device generative AI models progresses, mobile OS providers and app developers are increasingly developing apps, products, and services that incorporate features using these models. For example, various services—such as real-time call translation, audio file transcription, and photo background removal—are beginning to be offered. Since these functions are processed entirely on the device itself, they are gaining popularity among consumers who prioritize security and privacy.

Although these functions and services are still in the development stage at this point, it is expected that their scope of application and usage patterns will expand further as technology advances, leading to improved convenience for users. On the other hand, there is a possibility that competition policy issues may arise, such as when certain restrictive practices are implemented with respect to mobile operating systems (for discussion of the Antimonopoly Act and competition policy issues regarding restrictive practices concerning dedicated software on mobile operating systems, see Section 3.1 below).

⁴⁶ FTC “Partnerships Between Cloud Service Providers and AI Developers” January 17, 2025, p. 37 (<https://www.ftc.gov/news-events/news/press-releases/2025/01/ftc-issues-staff-report-ai-partnerships-investments-study>)

⁴⁷ For details on initiatives related to partnerships involving generative AI with overseas authorities, please refer to the attached document.

[Reference] Services that apply generative AI technology to existing fields of technology

- Market Overview in the Autonomous Driving Sector

I. Services that apply generative AI technology to existing fields of technology

As mentioned earlier, various services that directly utilize generative AI technology—such as ChatGPT and Copilot—have seen remarkable growth, and continue to evolve on a daily basis. However, the growth driven by generative AI technology is not limited to services that directly utilize such technology. In recent years, services that apply generative AI to existing fields of technology have also seen rapid growth. The application of generative AI to existing fields of technology is expected to drive innovation and significantly raise technical standards, and such efforts are already underway in areas such as autonomous driving and robotics. AI technologies, including generative AI used in fields such as autonomous driving and robotics, are sometimes referred to as “physical AI” because they involve physical operational outputs.

In the field of robotics, the development of AI robots capable of understanding human instructions in natural language and autonomously generating actions is underway through the application of generative AI technology and other innovations. In addition, AI robots are now capable of making complex judgments and performing human-like movements—tasks that were difficult to achieve with conventional technology—and efforts to replace certain tasks in manufacturing and logistics settings with AI robots are becoming increasingly widespread.

Furthermore, in the autonomous driving sector, the application of AI technologies—including generative AI—is enabling more flexible and context-aware driving decisions. Large-scale investment is expanding, particularly in the United States and China, and practical applications are advancing through companies such as Tesla in North America and Chinese startups. Within Japan as well, there is a rapidly growing trend across the entire automotive industry to adopt the latest technologies⁴⁸. On the other hand, the use of AI technologies—including generative AI—in the automotive industry, where safety is paramount, presents certain challenges, such as explainability⁴⁹, and the industry as a whole is working to address them.

In this way, the development and application of generative AI are expected to expand further through its application in existing fields of technology. We believe that, in order to understand the future competitive environment surrounding generative AI, it is important to look not only at fields that directly utilize generative AI technology, but also at these application fields.

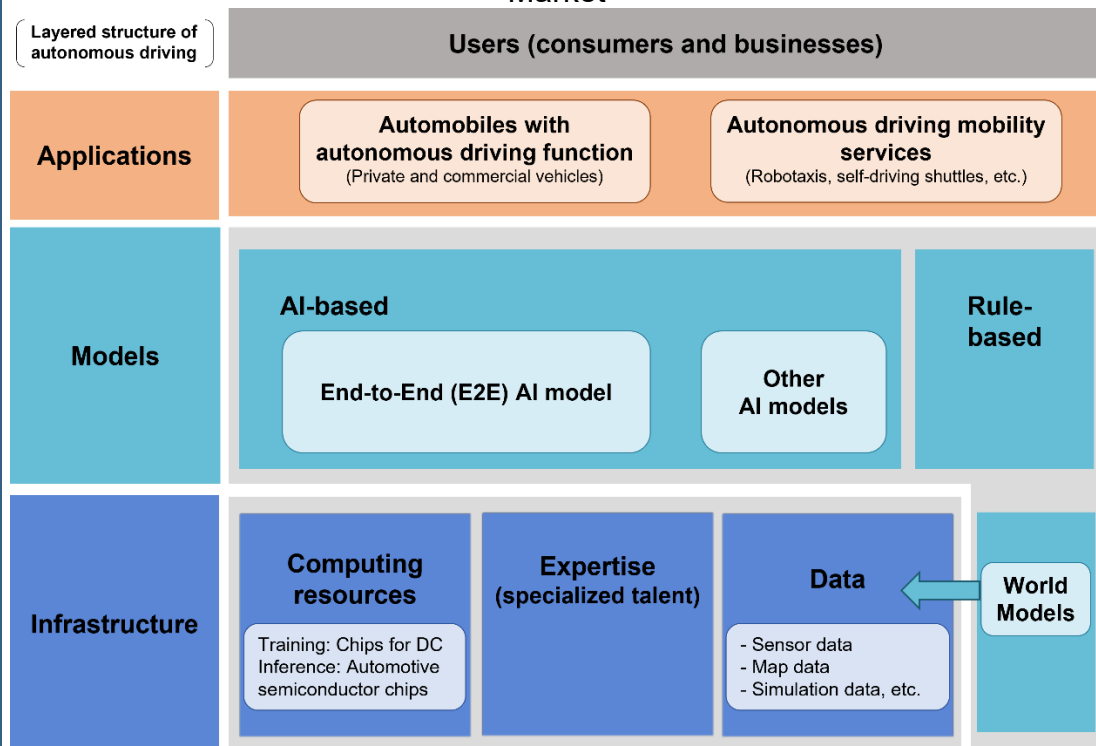
Therefore, in the following section, we will examine the current state of the autonomous driving sector—an area of the physical AI market that has garnered significant attention in recent years due to the application of AI technologies, including generative AI—based on information and opinions

gathered through hearings with domestic and international companies and other stakeholders.

II. Structure of the autonomous driving market

Similar to the generative AI market discussed earlier, the market structure of the autonomous driving sector consists of multiple layers, with domestic and overseas companies conducting business activities at each level. In this report, like the generative AI market, the autonomous driving market is categorized into three segments: “Infrastructure,” “Model,” and “Application,” as shown in Figure/Table 4.

Figure/Table 4. Diagram of Structure of the Autonomous Driving-related Market



Source: Prepared by the JFTC.

1. Infrastructure layer

(1) Computing resources

The technical characteristics of the semiconductor chips used in models for autonomous driving also differ between the training and inference phases. As described in Section 2.1(1) of the main text, the characteristics of each are as follows: during the training phase, semiconductor chips

⁴⁸ Nihon Keizai Shimbun “Cooperation among major automakers on AI autonomous driving to ensure safety, with the support of the government for domestic part production” May 29, 2025 (<https://www.nikkei.com/article/DGXZQQUA291U70Z20C25A5000000/>)

⁴⁹ The extent to which humans can understand and trust the reasoning behind decisions generated by machine learning algorithms

with high parallel processing capabilities are required to handle the enormous computational load; during the inference phase, semiconductor chips capable of high-speed, low-power processing are required to perform efficient inference. In particular, models used for autonomous driving must be capable of operating even in environments without a network connection; therefore, they are characterized by performing inference processing on onboard semiconductor chips (edge processing). As a result, automotive semiconductor chips are required to offer lower latency, lower power consumption, and greater resistance to vibration and heat compared to semiconductor chips used in the standard inference phase. Competition in the automotive semiconductor chip market is intense, with companies such as NVIDIA and Qualcomm, as well as firms like Mobileye—which focuses specifically on developing automotive semiconductor chips—actively engaged in this domain.

(2) Data

The data used to develop models for autonomous driving can be broadly categorized into three types: sensor data derived from vehicle driving data, map data, and simulation data generated in virtual environments. As noted in Section 2.1(2) of the main text, the quality of data is generally just as important as its quantity. However, since collecting such data requires significant resources—including sensor installation, ensuring driver quality, and data management—operators with a financial advantage and those that are already ahead in development are said to have an edge. On the other hand, some argue that since the type of data that needs to be collected varies depending on the specific functions a company aims to implement in autonomous driving, there is currently no significant difference in competitive advantage among operators.

Since autonomous driving requires safe operation under all traffic conditions, a key characteristic of the development of AI technology used in autonomous driving (hereinafter referred to as “autonomous driving AI”) is that the collection of data covering events with an extremely low probability of occurrence—so-called edge cases—as well as specific conditions such as rainy weather, has become a key competitive factor. Consequently, the use of synthetic data is gaining attention as a means of efficiently collecting such data. However, in the development of autonomous driving AI, there is a need to accurately reflect real-world driving conditions, and this presents challenges such as determining the appropriate amount of training data derived from simulated scenarios.

(3) Expertise (Specialized talent)

As noted in Section 2.1(3) of the main text, there is also a shortage of personnel specialized in the development of autonomous driving AI; in

particular, it has been pointed out that there is a limited number of individuals with advanced expertise in both generative AI and autonomous driving. Given this situation, competition for human resources has intensified, particularly in North America and China, and some have noted that it is difficult for domestic companies to secure specialized personnel, particularly in terms of compensation levels. On the other hand, some respondents noted that, thanks to factors such as brand recognition, they have been able to secure a certain number of local employees.

2. Model layer

Autonomous driving technology is shifting from traditional rule-based approaches—which rely on environmental perception using sensors (such as cameras, LiDAR, and radar) and high-precision 3D maps, and control vehicles based on programmed traffic and driving rules—toward AI-based methods centered on AI technologies, driven by the emergence of new technologies such as Transformer and large language models⁵⁰. The extent to which AI models are used in a given processing workflow, as well as which AI models are adopted, vary depending on the architecture; however, one AI model that has attracted particular attention is the E2E AI model⁵¹. Development is underway for AI-based autonomous driving models, such as end-to-end (E2E) AI models, including approaches where perception and path planning are handled by separate AI models (modular approach) and approaches where everything from perception to control is handled by a single AI model (monolithic approach)⁵². While North American and Chinese automakers are currently leading the way in the development of autonomous vehicles, each company is pursuing a different development strategy, and no single model or architecture has yet emerged as the clear frontrunner.

As mentioned earlier, while there are various AI models and architectures, development approaches can be broadly categorized into two types, and the autonomous driving AI used differs accordingly. The first approach involves offering affordable Level 2⁵³ autonomous vehicles—though not fully driverless—and using them as a foundation to improve autonomous driving performance over time. The second approach involves achieving Level 4 autonomous driving in specific urban areas and other limited

⁵⁰ Rule-based approaches have been central to the development of autonomous driving due to their affinity with legal regulations and traffic rules, but they have faced issues such as the vehicle coming to a halt in edge cases. In contrast, AI-based methods are said to be capable of handling edge cases because the AI, having learned human judgment by collecting vast amounts of driving data, can determine the optimal driving method.

⁵¹ A model with an architecture that uses AI to process perception, decision-making, and action.

⁵² The Ministry of Economy, Trade and Industry, the Ministry of Land, Infrastructure, Transport and Tourism “Mobility Digital Transformation (DX) Strategy 2025 Update” June 9, 2025, p. 30 (https://www.meti.go.jp/policy/mono_info_service/mono/automobile/jido_soko/pdf/mobilitydxsenryaku2025.pdf)

⁵³ Autonomous driving is classified into levels ranging from Level 0, where the driver is responsible for all operations, to Level 5, where the system handles all acceleration, steering, and braking at all times. Level 2 autonomous driving is defined as providing automated driving functions under specific conditions, while Level 4 autonomous driving is defined as providing fully automated driving under specific conditions. <https://www.mlit.go.jp/common/001226541.pdf>

environments, and then expanding the areas where fully driverless operation is possible. Generally speaking, in the former case, the proportion of AI-based autonomous driving models tends to be higher, as the focus is on achieving autonomous driving without limiting it to specific areas; in the latter case, rule-based and AI-based autonomous driving models tend to be used in combination, due to considerations of explainability and safety.

In addition, as mentioned earlier, synthetic data plays a crucial role in the development of autonomous driving, and there is intense global competition to develop world models⁵⁴ capable of generating the synthetic data used in the development of autonomous driving models.

Although development of autonomous driving models is currently led by companies in North America and China, other companies are conducting their own research and development—including on models and architectures—and there is intense competition in the field.

3. Application layer

In the autonomous driving sector, markets are emerging that provide end users with vehicles equipped with autonomous driving capabilities (such as passenger cars and commercial vehicles), as well as markets that offer autonomous mobility services such as robotaxis and autonomous shuttles.

As described above, while the autonomous driving sector shares many similarities with the market for services that directly utilize generative AI—such as the intense competition taking place at each layer—there are also many differences, including the nature of the training data and the types of models used.

Although no opinions were expressed during the current hearings that there are competition policy bottlenecks in the autonomous driving sector, as mentioned at the outset, it is important to consider these application fields as well in order to gain an understanding of the future competitive landscape surrounding generative AI.

⁵⁴ A model that learns the structure of the external world based on observational data obtained from it. It can simulate how a specific state changes and generate a video of it. <https://weblab.t.u-tokyo.ac.jp/news/20221130/>

Chapter 3. Antimonopoly Act and Competition Policy Issues Surrounding Generative AI

In the last report, we summarized our approach to “access restrictions and exclusion of competitors” and “tying” because comments received during the public comment period and hearings with businesses and other stakeholders raised competition concerns regarding these specific issues.

In this report, based on the state of the generative AI market summarized in Chapter 2 above, we have further examined the approaches outlined in the last report regarding the following issues—which are anticipated to be key issues from the perspectives of the Antimonopoly Act and competition policy: “restrictive practices concerning dedicated software on mobile operating systems” and “integrating generative AI into existing digital services,” and have subsequently reorganized our approach as follows.

1. Restrictive practices concerning dedicated software on mobile OS

As noted in Section 2.4(iv) above, recent years have seen an increase in the development of apps, products, and services that incorporate features (such as translation and summarization) utilizing on-device generative AI models. Apps, products, and services equipped with these features generally achieve the various functions described in Section 2.4(iv) above by accessing dedicated software (hereinafter referred to as “dedicated software”) provided by mobile OS providers on their mobile operating systems—typically via API connections—and by utilizing on-device generative AI models (hereinafter referred to as “native on-device generative AI models”) provided by such mobile OS providers. By accessing this dedicated software, app developers and other businesses will be able to utilize features that simplify the development of apps, products, and services, ensure security, and optimize performance. The following is a summary of the Antimonopoly Act’s approach to restrictive practices concerning dedicated software on mobile operating systems.

In February 2023, JFTC released a market study report on mobile operating systems⁵⁵. The report states that if a business operator holding a strong position in the mobile OS market restricts API access or takes similar measures against competing operators in the app market or related markets—thereby partially limiting access to smartphone functions via its own mobile OS (which can be accessed from its own apps, products, and services)—and thereby reduces business opportunities for or excludes such competing operators, this may constitute a violation of the Antimonopoly Act (such as private monopolization, unfair trade practices (paragraph 14 of the General Designation (Interference with a Competitor's Transactions), etc.))⁵⁶.

⁵⁵ Japan Fair Trade Commission “Market Study Report on Mobile OS and Mobile App Distribution” February 9, 2023 (<https://www.jftc.go.jp/en/pressreleases/yearly-2023/February/230209.html>)

⁵⁶ Japan Fair Trade Commission “Market Study Report on Mobile OS and Mobile App Distribution” February 9, 2023, p. 135-136 states the following views.

As mentioned above, apps, products, and services provided by mobile OS providers, app developers, and others access dedicated software in order to incorporate features that utilize native on-device generative AI models. In this context, there are concerns that mobile OS providers may restrict access to dedicated software—which can be accessed from their own apps, products, and services—for other app developers and businesses competing in the market for apps, products, and services that utilize native on-device generative AI models. The opinions of overseas model developers in response to such concerns are as follows:

[Opinions of Overseas Model Developers]

- Third-party applications can use the native on-device generative AI models just as our own applications do. Furthermore, even if both our application and a third-party application attempt to access the same relevant, dedicated software at the same time, our application will not be given priority simply because it is the genuine application. We firmly believe that an open ecosystem is key to fostering innovation, so we have no intention of creating any barriers that would disadvantage third parties, and we intend to adhere to this policy.

As shown in Figure/Table 5, if a business operator holding a strong position in the mobile OS market restricts access to dedicated software—which can be accessed from their own apps, products, and services—for other app developers and businesses competing in the market for apps, products, and services that utilize native on-device generative AI models, thereby reducing business opportunities for or excluding competing operators, such conduct may raise concerns under the Antimonopoly Act (private monopolization, unfair trade practices (paragraph 14 of the General Designation (Interference with a Competitor's Transactions), etc.)).

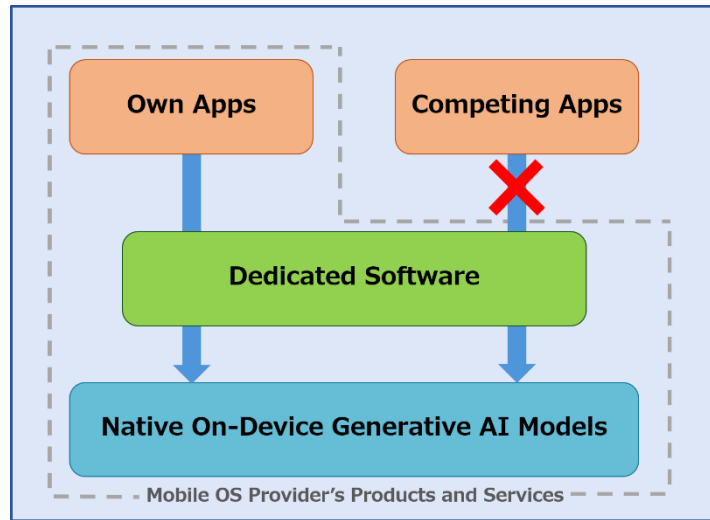
“When judging whether any conduct violates the AMA, various factors must be comprehensively considered. In assessing those security and privacy claims, consideration will be given to the rationality of the objective and the appropriateness of the means (whether there are alternative means that are less restrictive, etc.)
(Omitted)

The following is the evaluation of claims from the perspective of ensuring security and protecting privacy under the AMA.
- From the standpoint of safety, the removal of ineligible goods and services and ineligible enterprises is (omitted) the rationality of the objective is usually recognized when it can be assessed that an act was committed for the purpose of ensuring the security and protecting the privacy of consumers.

- On the other hand, regarding the appropriateness of the means, such as whether there is no other appropriate alternative to the conduct and whether they are socially appropriate means to achieve the objectives of ensuring the security and protecting the privacy of consumers, careful consideration should be given to whether there are other means that are less restrictive of competition in light of various specific circumstances.

- In determining the appropriateness of such means, it may be necessary to compare and evaluate individual technologies after receiving detailed technical information from the parties, and to quantitatively and qualitatively examine whether technologies other than those currently employed can be used to achieve the same or greater security and privacy protection.”

Figure/Table 5. Restrictive Practices concerning Dedicated Software on Mobile Operating Systems (1)



Furthermore, regardless of whether such actions are taken, the fact that on-device generative AI models accessible via dedicated software provided by mobile OS providers are limited to their own native on-device generative AI models may result in reduced business opportunities for on-device generative AI model developers and app developers other than mobile OS providers, and there are concerns about this situation. The opinions of overseas model developers in response to such concerns are as follows:

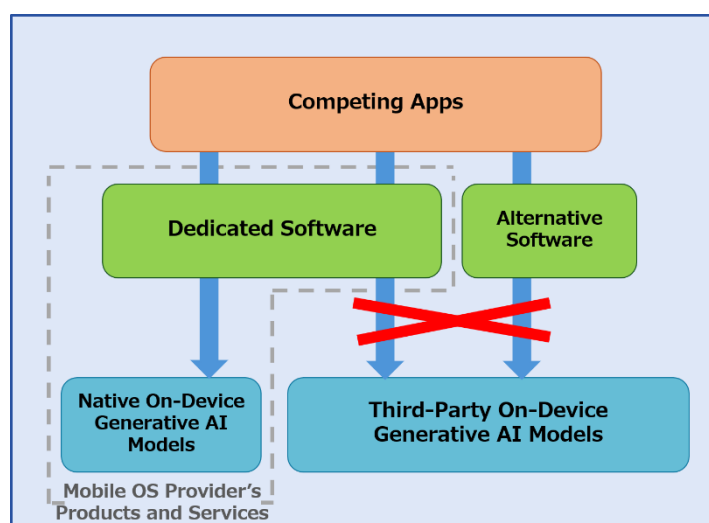
[Opinions of Overseas Model Developers]

- Because our mobile OS is inherently interoperable and open, third-party app developers can utilize a variety of open-source alternative tools to deploy third-party on-device generative AI models. In addition, we do not charge app developers any fees for running third-party on-device generative AI models on their devices. This means that third-party app developers will incur only the standard provisioning costs when running third-party on-device generative AI models on devices.

In the case of Figure/Table 6, it is important to determine whether on-device generative AI model developers and app developers can run third-party on-device generative AI models on devices by using software that replaces the dedicated software (hereinafter referred to as “alternative software”), in the same way that they would access the native on-device generative AI models through the dedicated software. For example, if a business operator with a strong position in the mobile OS market restricts access to on-device generative AI models—which can be accessed through dedicated software—to only that to native on-device generative AI models, other app developers and businesses competing in the market for apps, products, and services that use on-device generative AI models would be forced to develop solutions by combining multiple alternative software packages in order to run

third-party on-device generative AI models and if, due to reasons such as increased development costs, it becomes difficult to use third-party on-device generative AI models via such alternative software, such conduct may reduce business opportunities for other competing businesses or potentially exclude them, and could raise concerns under the Antimonopoly Act (private monopolization, unfair trade practices (paragraph 14 of the General Designation (Interference with a Competitor's Transactions), etc.)).

Figure/Table 6. Restrictive Practices concerning Dedicated Software on Mobile Operating Systems (2)



(Reference) Concept under the Mobile Software Competition Act (MSCA)

On June 12, 2024, the “Act on Promotion of Competition for Specified Smartphone Software” (Act No. 58 of 2024) was passed and enacted, and was promulgated on June 19 of the same year. As smartphones have rapidly become widespread and form the foundation of people’s daily lives and economic activities, this Act addresses software that is particularly essential for smartphone use (basic operation software, app stores, browsers, and search engines. These are collectively referred to as “specified software”). This Act imposes certain obligations on businesses designated by the JFTC (designated providers) as having a business scale related to the provision of specified software that meets or exceeds the threshold specified by Cabinet Order. The purpose is to establish a competitive environment that fosters innovation by diverse entities through competition, while ensuring security and other necessary measures, so that consumers can choose from the diverse services resulting from such innovation and enjoy the benefits thereof. This Act came into full effect on December 18, 2025.

Article 7, Item 2 of the Act stipulates that a prohibited act by a designated provider with respect to basic operation software is any act that prevents other providers from using functions related to the operation of smartphones

controlled by such basic operation software (hereinafter referred to as “OS functions”)—which the designated provider (including its subsidiaries, etc.; the same applies hereinafter) uses to provide individual software—for the provision of individual software with equivalent performance.

Therefore, if the aforementioned restrictive practices concerning dedicated software on the mobile OS constitute acts that prevent other providers from using OS functions—which the designated provider associated with the basic operation software utilizes to provide individual software—with equivalent performance, such acts fall into the category listed in Article 7, Item 2 of the Act (interference with the use of OS functions), and, unless there are justifiable grounds for such acts, they would constitute a violation of the provisions of that article.

Furthermore, even if the requirements for a prohibited act under Article 7, Item 2 of the Act are not formally met, if the conduct constitutes unfair treatment of an individual app provider, it will constitute a violation of Article 6 of the Act.

2. Integrating generative AI into existing digital services

Generally speaking, combining multiple functions to provide services to customers that add new value is a common approach to technological innovation and sales promotion; the integration of existing digital services with generative AI does not, in and of itself, immediately raise concerns regarding the Antimonopoly Act.

However, there are concerns that if a business operator with a strong position in a specific digital services market integrates generative AI into that digital service and offers it to users, this could create conditions that hinder the business activities of existing competitors in the generative AI model or generative AI application markets, or raise barriers to entry, and concerns have been voiced regarding such a situation. The opinions of overseas model developers in response to such concerns are as follows:

[Opinions of Overseas Model Developers]

- Currently, generative AI functionality is offered in all applications, etc. and it is indispensable requirement for application providers to use and make available generative AI functionality into their digital services in order to improve the products for customers and remain competitive. Accordingly, we are offering generative AI functionality in our existing digital services in order to improve the experience for users through direct access to generative AI.
- In general, integrating generative AI into existing digital services allows users to reap benefits such as increased productivity and creativity. For example, we believe that integrating our company’s generative AI into our services will deliver significant benefits in terms of productivity. For instance, this will enable the AI assistant to access the necessary context (background information) related to a work document without the user having to provide a detailed explanation in the prompt (question or task instruction).

Integrating generative AI into existing digital services can involve providing document creation applications equipped with generative AI assistant features, or offering office suites that incorporate generative AI chatbots. For such conduct to be considered a tie-in sale, the generative AI (the tied product) integrated into an existing digital service (the tying product or service⁵⁷) must qualify as “other products.”

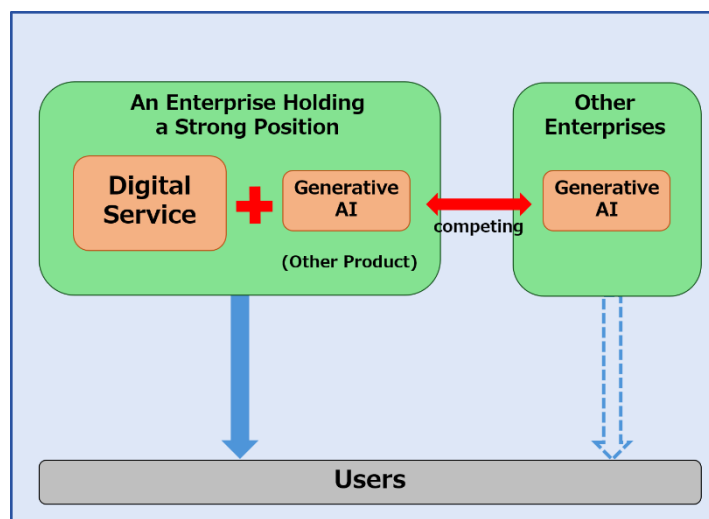
Whether a product sold in conjunction with one product falls under “other products” is determined by whether the combined products each possess distinct characteristics and are traded independently. Specifically, for each product, factors such as “whether the target consumers differ,” “whether the content or functions differ (including whether the content or functions of the combined product have substantially changed compared to those of the individual products prior to the bundling),” and “whether target consumers can purchase the products individually (including whether the combined product is typically sold or used as a single unit)” are taken into comprehensive consideration⁵⁸.

Based on this reasoning, as illustrated in Figure/Table 7, when a tied, subordinate generative AI falls under “other products,” if the integration of such an existing digital service with generative AI is carried out by a provider holding a strong position in the relevant existing digital service market, and this results in a market foreclosure effect on the market of the tied product, generative AI—that is, for providers providing generative AI models or applications, or those seeking to begin providing such models or applications, it becomes difficult to secure users; that is, customers, the costs required for business activities are increased, and the incentive for new market entry or new product development is undermined—thereby creating a situation where existing competitors or new entrants are excluded or their business opportunities are reduced—such integration constitutes a tie-in sale and may raise issues under the Antimonopoly Act (private monopolization, unfair trade practices (paragraph 10 of the General Designation (Tie-in Sales, etc.))).

⁵⁷ Hereinafter, the terms “products” and “services” are simply referred to as “products” without any particular distinction.

⁵⁸ “The Guidelines for Exclusionary Private Monopolization under the Antimonopoly Act,” Section 2-4(1); “Guidelines concerning Distribution Systems and Business Practices under the Antimonopoly Act,” Part 1, Section 2-7(3)

Figure/Table 7. Integrating Generative AI into Existing Digital Services



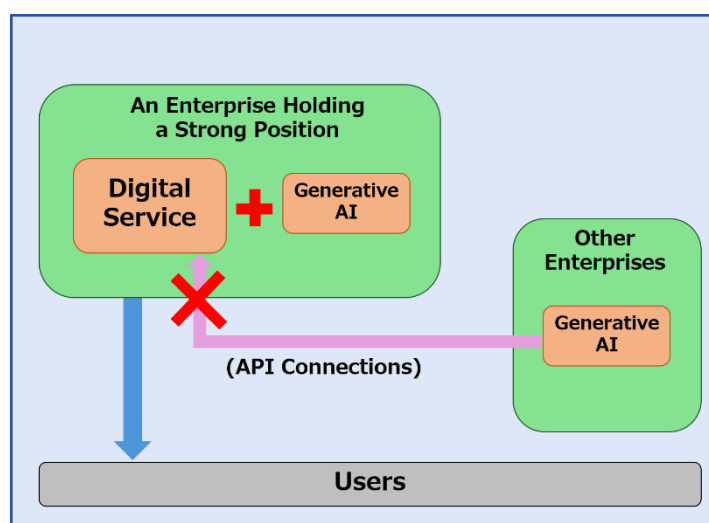
Furthermore, if a provider in a specific digital services market integrates its own generative AI with that digital service, it is conceivable that the provider might restrict API access to that digital service for generative AI provided by competing generative AI model developers. The opinions of overseas model developers in response to such concerns are as follows:

[Opinions of Overseas Model Developers]

- Our company provides third parties with fully open interoperability capabilities. This interoperability mechanism enables third parties to build solutions (including AI solutions) that effectively interoperate with our applications and solutions. By ensuring this level of broad interoperability, we have been able to offer users of our digital services a wide range of options.
- We allow AI assistant features powered by competing generative AI models to be offered on our platform and to access our platform. We believe that doing so will foster innovation within our platform's ecosystem.

For example, as shown in Figure/Table 8, if a provider holding a strong position in a specific digital services market prevents users from using competing generative AI models within that digital service—by restricting API access to the service or taking similar actions for generative AI provided by competing generative AI model developers—without a reasonable justification such as a technical necessity, or beyond the scope of such necessity, this may raise issues under the Antimonopoly Act (private monopolization, unfair trade practices (paragraph 14 of the General Designation (Interference with a Competitor's Transactions), etc.)).

Figure/Table 8. Acts of Restricting API Connections, etc. to Existing Digital Services



(Reference) Acts of providing a combination of cloud services and generative AI

With the development of the generative AI market, the cloud services market is experiencing rapid growth, as evidenced by increased revenue at three major cloud service providers. Although no specific concerns were raised during the current hearings, if, for example, a company with a strong position in the cloud services market were to integrate its own generative AI into its cloud services and offer them as a package, thereby creating a foreclosure effect in the generative AI market, such integration could constitute a tie-in sale and raise issues under the Antimonopoly Act (private monopolization; unfair trade practices (paragraph 10 of the General Designation (Tie-in Sales, etc.))). Furthermore, if a provider holding a strong position in the generative AI market sets unreasonably high license fees for its generative AI exclusively for users of cloud services provided by its competitors, thereby creating a foreclosure effect in the cloud services market, this may constitute discriminatory pricing, discriminatory treatment regarding transaction conditions, or interference with a competitor's transactions, and could raise issues under the Antimonopoly Act (private monopolization; unfair trade practices (paragraph 3 of the General Designation (Discriminatory Consideration), paragraph 4 of the General Designation (Discriminatory Treatment on Trade Terms, etc.), and paragraph 14 of the General Designation (Interference with a Competitor's Transactions))). The JFTC will continue to monitor the cloud services market.

In June 2022, the JFTC published a "Report on Trade Practices in Cloud Services Sector." This report outlines the legal principles under the Antimonopoly Act regarding practices that could adversely affect competition in the cloud services market, such as the integration of distinct, independently traded functions within the market and preferential treatment in the licensing

of software used in cloud services.

Subsequently, the JFTC initiated an investigation into alleged violations of the Antimonopoly Act⁵⁹ by Microsoft Corporation, Microsoft Japan Co., Ltd., and Microsoft Ireland Operations Limited (hereinafter collectively referred to as “Microsoft Corporation, etc.”) regarding their conduct in the cloud services market. As part of this investigation, it decided on March 4, 2026, to seek information and opinions from third parties.

3. Other points of discussion

In the discussion paper released at the start of this survey, we provided examples of “self-preferencing,” “parallel conduct using generative AI,” and “acquisition of highly skilled professionals through partnerships” in addition to the two main discussion points, with the aim of contributing to future discussions.

The summary of these three points presented in the discussion paper is as follows.

(1) Self-preferencing

If a provider of generative AI models develops such a model to favor its own products or services over another product in its inference results, this could potentially affect competition regarding those products or services.

(2) Parallel conduct using generative AI

While price competition may intensify due to price surveys and other activities conducted using generative AI, it is anticipated that the use of common underlying data and algorithms could lead to identical or similar pricing strategies and production targets, which could potentially affect competition.

(3) Acquisition of highly skilled professionals through partnerships

If the purpose of forming a partnership is to secure highly skilled professionals and the partnership effectively results in a business transfer, it may have an impact on competition.

⁵⁹ Microsoft Corporation, etc., which provides cloud service known as Microsoft Azure (hereinafter referred to as “Azure”), is suspected to prevent enterprises which provide their own cloud services competing with Azure (hereinafter referred to as “Competing Cloud Services”) from obtaining transactions related to provision of cloud services, by engaging in the following conduct, in relation to use of cloud services, towards enterprises that have actually obtained or may potentially obtain license to use Microsoft’s software or services, named as Windows Server, Windows Client, Microsoft SQL Server, Microsoft 365, Visual Studio, etc. (hereinafter referred to as the “Services”):

(1) not allowing them to use the Services in combination with Competing Cloud Services; or
(2) modifying or setting trade terms of the Services so that costs for the users, etc. are increased when the Services are used in combination with Competing Cloud Services compared to the case in which the Services are used on Azure.

To date, we have not received any comments describing conduct similar to the examples provided regarding these three points. As noted above, this report further examines the concepts of “access restrictions and exclusion of competitors” and “tying”—which were outlined in the last report—and has redefined them as “restrictive practices concerning dedicated software on mobile operating systems” and “the integration of generative AI into existing digital services.” Moving forward, it will be necessary to continue monitoring the market, including consideration of these three points.

Chapter 4. Next steps for the Japan Fair Trade Commission

The current market size for generative AI in Japan is estimated at 665.3 billion yen (forecast for 2025). It is projected to grow rapidly at an average annual rate of 38.1%, reaching 1.9791 trillion yen by 2029⁶⁰, and the market is expected to continue expanding and growing. Against this backdrop, in Japan, the “Bill on Promotion of Research and Development, and Utilization of AI-related Technology” was submitted to the 217th Regular Session of the Diet on February 28, 2025, and was passed and enacted at the plenary session of the House of Councilors on May 28 of the same year. In addition, pursuant to the Act, the “Artificial Intelligence Basic Plan” was approved by the Cabinet on December 23 of the same year. In addition, as shown in the attached document, several competition authorities in various countries and regions overseas have taken steps such as publishing reports on their fact-finding investigations. As the generative AI market continues to grow, efforts are being made both domestically and internationally to actively address the challenges associated with generative AI.

This report compiles the findings from our previous investigations, provides an updated overview of the current state of the generative AI market, and examines the relevant Antimonopoly Act issues in greater detail. The JFTC hopes that the views set out in this report will help prevent potential adverse effects in the generative AI market and further promote fair and free competition. Furthermore, we will continue to respond strictly and appropriately to any specific cases that raise concerns regarding the Antimonopoly Act, including not only the conduct identified in this report as problematic under the Antimonopoly Act, but also conduct raised as issues in Chapter 3 without specific examples being cited, as well as conduct not raised as issues at all.

With regard to potential adverse effects in the generative AI market, we will actively work in close coordination and cooperation with relevant government agencies as necessary to ensure a fair competitive environment. Furthermore, since some of the competitors in the generative AI market operate on a global scale, competition authorities in various countries and regions around the world are taking a keen interest in—and expressing significant concerns about—the business activities of such providers. The JFTC will continue to exchange views with competition authorities in various countries and regions at multiple levels, and will work in close cooperation with relevant overseas authorities—utilizing forums such as the International Competition Network (ICN) and the Organisation for Economic Co-operation and Development (OECD)—to foster a competitive environment.

The generative AI market environment is subject to daily changes and technological innovations that can impact its structure, and as it is likely to undergo significant changes in the future, it is crucial to accurately assess market trends. We hope that the generative AI market will continue to

⁶⁰ Fuji Chimera Research Institute, Inc. “2026 Comprehensive Survey of the AI Market: Soaring Growth through Generative AI and AI Agents, Market Edition” January 20, 2026, p. 5.

develop under conditions of fair and free competition.

(Appendix) Foreign Competition Authorities' Work in Generative AI Space

1 EU

On January 9, 2024, the European Commission launched a call for opinions on competition in the virtual worlds and generative AI sectors (with a deadline of March 11 of that year)⁶¹, and on June 28 of the same year, it held a workshop on the theme of “competition in the context of virtual worlds and generative AI” as a follow-up to the consultation⁶². The workshop also revealed that, while it had conducted an investigation into the Microsoft-OpenAI partnership from the perspective of EU merger regulations, it concluded that Microsoft had not acquired permanent control over OpenAI and had therefore closed the investigation⁶³.

On September 19 of that year, based on the comments submitted during the call for opinions, insights gained from the workshop, and the European Commission’s investigation, the Commission published a policy brief on competition in generative AI and virtual worlds⁶⁴. The policy brief identifies trends in the generative AI market, such as vertical integration and the expansion of partnerships, the miniaturization of models and improvements in efficiency, and the coexistence of open-source and closed-source models, while also highlighting competitive concerns such as restrictions on access to key resources, exclusionary practices leveraging market dominance, and network effects. The Commission says it will continue to address potential issues in the generative AI sector in order to ensure a level playing field.

On February 9, 2026, the European Commission announced that it had issued a statement of objections to Meta, stating that the company’s exclusion of competitors’ AI chatbots from the WhatsApp platform—based on the terms of service for “WhatsApp Business Solutions” that took effect on October 15, 2025—may violate European competition law⁶⁵. In addition, on December 9, 2025, the Commission announced that it had begun reviewing Google’s use of web content for AI purposes⁶⁶. The review examines whether the company is distorting competition by imposing terms of service that are disadvantageous to publishers and content creators, and whether it is putting competing AI model developers at a disadvantage by holding privileged access to such content.

Furthermore, on January 27, 2026, the Commission initiated specific proceedings under the Digital Markets Act (DMA), seeking to ensure interoperability—specifically, that third parties be granted equivalent access to the hardware and software features used by Google’s AI services, such as its AI Assistant Gemini, on the Android OS—and to require the sharing of anonymized search data, including rankings, queries, and clicks, with third

⁶¹ https://ec.europa.eu/commission/presscorner/detail/en/ip_24_85

⁶² https://competition-policy.ec.europa.eu/about/reaching-out/virtual-worlds-and-generative-ai_en

⁶³ https://ec.europa.eu/commission/presscorner/detail/en/speech_24_3550

⁶⁴ <https://digital-strategy.ec.europa.eu/en/news/commission-publishes-policy-brief-competition-generative-ai-and-virtual-worlds>

⁶⁵ https://ec.europa.eu/commission/presscorner/detail/en/ip_26_310

⁶⁶ https://ec.europa.eu/commission/presscorner/detail/en/ip_25_2964

parties⁶⁷.

2. United Kingdom

The CMA (Competition and Markets Authority) conducted an investigation into AI foundation models from the perspective of competition and consumer protection policies. It published a report on generative AI foundation models on September 18, 2023⁶⁸, and released an updated version of the report on April 11, 2024⁶⁹. The report indicated that there are concerns regarding fairness, efficiency, and openness in competition.

In addition, the CMA had been conducting merger reviews regarding partnerships between several Big Tech companies and generative AI startups, and has decided that none of these partnerships would proceed to the second phase of the review.

- Microsoft and Mistral AI (review closed on May 17, 2024)⁷⁰
- Microsoft and Inflection AI (review closed on September 4 of the same year)⁷¹
- Amazon and Anthropic (review closed on September 27)⁷²
- Google and Anthropic (review closed on November 19 of the same year)⁷³
- Microsoft and OpenAI (review closed on March 5, 2025)⁷⁴

On October 10, 2025, the CMA announced that it had designated Google as an operator with a strategic market status (SMS)⁷⁵ in the search and search advertising sectors under the Digital Markets, Competition and Consumers Act (DMCCA)⁷⁶. While stating that Gemini, Google's AI Assistant, is currently exempt from regulation, the authority plans to subject AI-based search features—such as AI overview and AI mode—to regulation.

3. United States

Engineers from the FTC (Federal Trade Commission)'s Office of Technology provide insights on technology-closely-related competition concerns in the Technology Blog⁷⁷. In a June 29, 2023, post titled "Generative AI Raises Competition Concerns," the blog explains that if the elements essential to providing generative AI are controlled by specific providers, this could harm

⁶⁷ https://ec.europa.eu/commission/presscorner/detail/en/ip_26_202

⁶⁸ https://assets.publishing.service.gov.uk/media/650449e86771b90014fdab4c/Full_Non-Confidential_Report_PDFa.pdf

⁶⁹ <https://www.gov.uk/government/publications/ai-foundation-models-update-paper>

⁷⁰ <https://www.gov.uk/cma-cases/microsoft-slash-mistral-ai-partnership-merger-inquiry>

⁷¹ <https://www.gov.uk/cma-cases/microsoft-slash-inflection-ai-inquiry#invitation-to-comment-closed>

⁷² <https://www.gov.uk/cma-cases/amazon-slash-anthropic-partnership-merger-inquiry>

⁷³ <https://www.gov.uk/cma-cases/alphabet-inc-google-llc-slash-anthropic-merger-inquiry>

⁷⁴ <https://www.gov.uk/cma-cases/microsoft-slash-openai-partnership-merger-inquiry>

⁷⁵ The CMA may designate providers that hold a strategic market status in specific digital sectors; once a designation is made, the CMA may impose behavioral requirements on the designated provider or introduce pro-competitive interventions designed to deliver outcomes that are beneficial to UK consumers and businesses.

⁷⁶ <https://www.gov.uk/government/news/cma-confirms-google-has-strategic-market-status-in-search-services>

⁷⁷ <https://www.ftc.gov/policy/advocacy-research/tech-at-ftc>

competition⁷⁸. The report also summarizes the key points raised during the various panels at the AI Tech Summit⁷⁹ held by the FTC on January 25, 2024⁸⁰.

On January 25, 2024, the FTC issued an order under Article 6 of the FTC Act requiring Alphabet, Amazon, Anthropic, Microsoft, and OpenAI to provide information regarding recent investments and partnerships involving generative AI companies and major cloud service providers⁸¹, and on January 17, 2025, it published a staff report titled “AI Partnerships and Investments⁸².” The report raises concerns that partnerships between large-scale cloud service providers (Alphabet, Amazon, and Microsoft) and generative AI developers (Anthropic and OpenAI) could give rise to competition policy issues, such as restrictions on access to input data, preferential access, increased switching costs, and the acquisition of confidential information⁸³.

4. EU, UK, US

On July 23, 2024, the European Commission, the CMA, the DOJ (Department of Justice Antitrust Division), and the FTC issued a joint declaration on competition in the generative AI foundation model and AI product markets⁸⁴. The key points are as follows:

- Although enforcement is carried out at the discretion of each national competition authority, given that risks related to AI are likely to manifest across borders, the authorities will strive to share a proper understanding of the issues and, where necessary, exercise their respective powers.

⁷⁸ <https://www.ftc.gov/policy/advocacy-research/tech-at-ftc/2023/06/generative-ai-raises-competition-concerns>

⁷⁹ <https://www.ftc.gov/news-events/events/2024/01/ftc-tech-summit>

At the FTC’s AI Tech Summit, we brought together participants from academia, industry, civil society organizations, and government agencies to hold three panel discussions focusing on different layers of AI: hardware and infrastructure, data and models, and consumer applications.

⁸⁰ “Semiconductor Chips & Cloud Computing” March 14, 2024:

<https://www.ftc.gov/policy/advocacy-research/tech-at-ftc/2024/03/semiconductor-chips-cloud-computing-quote-book>

“Data and Models” April 17, 2024:

<https://www.ftc.gov/policy/advocacy-research/tech-at-ftc/2024/04/data-models-quote-book-tech-summit-ai>

“Consumer Facing Applications” April 24, 2024:

<https://www.ftc.gov/policy/advocacy-research/tech-at-ftc/2024/04/consumer-facing-applications-quote-book-tech-summit-ai>

⁸¹ <https://www.ftc.gov/news-events/news/press-releases/2024/01/ftc-launches-inquiry-generative-ai-investments-partnerships>

⁸² <https://www.ftc.gov/news-events/news/press-releases/2025/01/ftc-issues-staff-report-ai-partnerships-investments-study>

⁸³ It should be noted that Chairman Andrew N. Ferguson (who was a member at the time the report was published) and Member Melissa Holyoak (at the time) expressed their support for the publication of the report itself, but stated that Chapter 5 (which highlights areas of concern regarding the potential impacts of AI partnerships) should be skipped or read with a skeptical eye.

<https://www.ftc.gov/legal-library/browse/cases-proceedings/public-statements/concurring-dissenting-statement-commissioner-andrew-n-ferguson-joined-commissioner-melissa-holyoak>

⁸⁴ https://competition-policy.ec.europa.eu/about/news/joint-statement-competition-generative-ai-foundation-models-and-ai-products-2024-07-23_en

<https://www.gov.uk/government/publications/joint-statement-on-competition-in-generative-ai-foundation-models-and-ai-products/joint-statement-on-competition-in-generative-ai-foundation-models-and-ai-products>

<https://www.justice.gov/opa/pr/leaders-justice-department-federal-trade-commission-european-commission-and-uk-competition>

<https://www.ftc.gov/news-events/news/press-releases/2024/07/ftc-doj-international-enforcers-issue-joint-statement-ai-competition-issues>

- Potential threats to competition in the AI market include the possibility that a small number of companies could exploit bottlenecks across existing or emerging AI-related fields to gain disproportionate influence over future developments—such as specialized chips and massive computing power for developing foundation models—or that existing large corporations could expand or consolidate the positions they have established through past major technological transformations, thereby hindering future competition.
- To foster competition and drive innovation within the AI ecosystem, fair trade practices, interoperability, and other factors are essential.

5. France

On February 8, 2024, the French Competition Authority launched a call for opinions on generative AI (with a deadline of March 22 of that year)⁸⁵, and published a report on June 28 of the same year⁸⁶. The report discusses the high barriers to entry in the generative AI market and how generative AI enhances competitive advantages in other digital sectors. Furthermore, after explaining that there are competitive concerns in the upstream segment of the value chain, recommendations are made to address these concerns.

Additionally, on December 17, 2025, the French Competition Authority published the findings of its investigation into competition issues related to the impact of AI on energy and the environment⁸⁷, calling for measures to prevent large corporations from monopolizing resources and to ensure both sustainability and a healthy competitive environment. Subsequently, on January 9, 2026, the authority announced that it would begin a fact-finding survey on conversational agents such as ChatGPT and Gemini⁸⁸, and on the 29th of the same month, it launched a call for opinions (with a deadline of March 6 of the same year)⁸⁹.

In addition, on September 27, 2023, the French Competition Authority conducted an investigation, with court authorization, into companies suspected of engaging in anti-competitive practices in the graphics card sector⁹⁰.

6. Canada

⁸⁵ <https://www.autoritedelaconurrence.fr/en/press-release/generative-artificial-intelligence-autorite-starts-inquiries-ex-officio-and-launches>

⁸⁶ https://www.autoritedelaconurrence.fr/en/press-release/generative-artificial-intelligence-autorite-issues-its-opinion-competitive?referrer=content_seehereview

⁸⁷ <https://www.autoritedelaconurrence.fr/en/press-release/autorite-publishes-its-study-competition-issues-surrounding-energy-and-environmental>

⁸⁸ <https://www.autoritedelaconurrence.fr/en/press-release/conversational-agents-autorite-starts-inquiries-ex-officio-view-issuing-opinion>

⁸⁹ <https://www.autoritedelaconurrence.fr/en/article/autorite-lance-public-consultation-conversational-agents>

⁹⁰ <https://www.autoritedelaconurrence.fr/en/press-release/general-rapporteur-autorite-de-la-concurrence-indicates-unannounced-inspection-was>

It should be noted that while the French Competition Authority has announced that it is conducting an investigation into this matter, it has not disclosed the identity of the entity under investigation.

On March 20, 2024, the Competition Bureau Canada published a discussion paper on competition in the AI sector⁹¹ to deepen understanding and foster discussion regarding how competition in AI-related markets is evolving, how competition in these markets can be protected and promoted, and how to address competition concerns. The Bureau also included questions in the discussion paper to solicit public comments (with a deadline of July 7 of the same year), and 28 submissions were received⁹². On January 27, 2025, the government released a report based on public comments⁹³, noting concerns that large-scale investments by Big Tech companies and others could hinder competition and innovation, while also stating that it remains unclear whether current antitrust laws can effectively address the anti-competitive behavior that AI may facilitate. The Bureau stated that it would continue to engage with stakeholders to deepen its understanding of new competition-related issues surrounding AI.

On September 16, 2024, the Canada's Competition Summit 2024⁹⁴ was held under the theme "Market Dynamics in the AI Era," and a report⁹⁵ summarizing the summit was published on November 7 of the same year.

7. Korea

On August 1, 2024, the KFTC (Korea Fair Trade Commission) announced that it had launched a written survey of major domestic and international companies in the AI industry to identify potential competition and consumer rights issues⁹⁶. The report was published on December 17 of the same year⁹⁷. The significance of the report lies in its overview of the competitive landscape of the rapidly evolving domestic generative AI market—which is still in its early stages—across the entire value chain. While the generative AI market is a dynamic one characterized not only by vigorous competition but also by rapid change and development driven by technological innovation, concerns have emerged regarding the market following the emergence of companies with dominant market power in specific segments. Consequently, the agency plans to continue monitoring the competitive situation in the generative AI market and conduct additional market analysis as necessary.

⁹¹ <https://www.canada.ca/en/competition-bureau/news/2024/03/competition-bureau-seeks-feedback-on-artificial-intelligence-and-competition.html>

⁹² <https://competition-bureau.canada.ca/en/how-we-foster-competition/consultations/written-responses-consultation-artificial-intelligence-and-competition-discussion-paper>

⁹³ <https://www.canada.ca/en/competition-bureau/news/2025/01/competition-bureau-issues-report-summarizing-feedback-on-artificial-intelligence-and-competition.html>

⁹⁴ <https://www.canada.ca/en/competition-bureau/news/2024/09/canadas-competition-summit-2024-competition-bureau-releases-details-about-panels-and-expert-participants.html>

⁹⁵ <https://competition-bureau.canada.ca/en/how-we-foster-competition/education-and-outreach/report-summit-2024-competition-ai-era>

⁹⁶ https://www.ftc.go.kr/www/selectBbsNttView.do?pageUnit=10&pageIdx=32&searchCnd=all&key=12&bordCd=3&searchCtgr=01_02&nttSn=43530

⁹⁷ https://www.ftc.go.kr/www/selectBbsNttView.do?pageUnit=10&pageIdx=14&searchCnd=all&key=12&bordCd=3&searchCtgr=01_02&nttSn=43730

8. Taiwan

On July 18, 2025, the TFTC (Taiwan Fair Trade Commission) published a “Public Consultation Paper for Generative Artificial Intelligence and Competition” to gain a better understanding of the supply chain and competitive landscape in the AI-related industry, and solicited public comments (the deadline for submissions was September 7 of the same year)⁹⁸. On March 18, 2026, the TFTC published a “Report and Policy Statement”⁹⁹ based on public comments, stating that it would adjust its enforcement strategy in a timely and prudent manner and take appropriate measures to address competition concerns in the rapidly evolving generative AI market.

9. Italy

On July 30, 2025, the AGCM (Italian Competition and Market Authority) launched an investigation into Meta on suspicion of violating Article 102 of the Treaty on the Functioning of the European Union (abuse of a dominant position)¹⁰⁰. The AGCM took issue with Meta’s actions, which involved leveraging WhatsApp’s dominant market position to pre-install Meta AI on WhatsApp, place it in a prominent location within the app, and integrate it into the search bar. In addition, based on the terms of service for “WhatsApp Business Solutions” that Meta implemented on October 15 of that year, the AGCM announced on November 26 of that year that it would also include in its investigation Meta’s actions to exclude competitors’ AI chatbots from the WhatsApp platform¹⁰¹. On December 24 of that year, the court issued a preliminary injunction ordering WhatsApp to suspend its policy of excluding rival AI chatbots from its platform, on the grounds that the policy could distort competition in the Italian AI chatbot market¹⁰².

10. Portugal

On November 6, 2023, the Portuguese Competition Authority published an analysis on generative AI and competition, along with an issue paper, to raise awareness of the competitive risks in the generative AI sector¹⁰³. Subsequently, the authority published a short paper on data¹⁰⁴ in September 2024, a short paper on models¹⁰⁵ in December of the same year, a short paper on securing human resources in the generative AI field¹⁰⁶ in July 2025,

⁹⁸ <https://www.ftc.gov.tw/internet/english/doc/docDetail.aspx?docid=18159&uid=179>

⁹⁹ <https://www.ftc.gov.tw/internet/english/doc/docDetail.aspx?docid=18369&uid=179>

¹⁰⁰ <https://en.agcm.it/en/media/press-releases/2025/7/A576>

¹⁰¹ <https://en.agcm.it/en/media/press-releases/2025/11/A576>

¹⁰² <https://en.agcm.it/en/media/press-releases/2025/12/A576>

¹⁰³ <https://www.concorrenza.pt/en/articles/adc-warns-competition-risks-generative-artificial-intelligence-sector>

¹⁰⁴ <https://www.concorrenza.pt/en/articles/adc-warns-competition-risks-regarding-access-and-use-data-generative-ai>

¹⁰⁵ <https://www.concorrenza.pt/en/articles/adc-warns-competition-risks-regarding-access-generative-ai-models>

¹⁰⁶ <https://www.concorrenza.pt/en/articles/adc-highlights-talent-shortages-and-restrictions-labour-mobility-hamper-competition-ai>

and a short paper on access to chips¹⁰⁷ in February 2026.

11. G7 Competition Summit

On October 3 and 4, 2024, the AGCM hosted the “G7 Competition Summit” in Rome, Italy, attended by the heads of G7 competition authorities and policy-making bodies, among others; the JFTC also attended.

At the summit, discussions were held on topics such as competition concerns within the AI value chain (key inputs and partnerships in AI), competition-related issues concerning AI (downstream and adjacent markets for AI, as well as collusion involving algorithms), and regulatory and policy approaches to the AI market. As a result, the “Digital Competition Communique” was adopted as the summit’s outcome document. The Communique indicated the approach to:

- Creating innovation with AI
- Competitive concerns regarding AI
- Guiding principles for promoting competition and innovation
- Roles of competition authorities (Strict enforcement of the law, strengthening international cooperation, etc.), etc.¹⁰⁸

¹⁰⁷ <https://www.concorrenza.pt/en/articles/adc-highlights-potential-competition-risks-artificial-intelligence-chip-value-chain>

¹⁰⁸ <https://www.jftc.go.jp/en/pressreleases/yearly-2024/October/1007.html>