Event Report & Q&A

Dealing with the Brussels Effect: How should Japanese companies prepare for the EU-AI Act? 4

Date	Monday, 12 May 2025, 17:30-18:30 JST
Venue	ZOOM Webinar
Organized by	Tokyo College, The University of Tokyo; Institute for Future Initiatives, The University of Tokyo; Next Generation Artificial Intelligence Research Center, The University of Tokyo
Supported by	Japan AI Safety Institute; Osaka University Research Center on Ethical, Legal and Social Issues; Japan Deep Learning Association

Introduction

The EU Artificial Intelligence Act (AI Act), approved by the EU Council on May 21, 2024, and entered into force on August 1, 2024, is the world's first comprehensive regulatory framework for artificial intelligence (AI)[1]. Its provisions will be implemented in stages by December 31, 2030, with potential implications for companies and organizations outside the EU, including those in Japan.

The University of Tokyo hosted a webinar event titled "Dealing with the Brussels Effect: How should Japanese companies prepare for the EU-AI Act?"[2] on December 11, 2024, where experts provided insights into the AI Act and the Code of Practice (CoP) for general-purpose AI (GPAI), which is currently in the drafting process and will detail the AI Act rules on GPAI. At the second event[3] held on January 15, 2025, the second draft was reviewed and discussed, followed by the third draft at the third event[4] on March 19. This fourth event was organized in response to the public consultation[5] launched by the EU-AI Office, featuring overview of the public consultation and key points for Japanese companies to note.

This Event Report provides a summary of the event proceedings, along with a selection of questions from attendees and the speakers' responses to them.

[1] The AI Act is a regulation under EU law, which means it is directly applicable in all EU member states.

[2] First event: https://www.tc.u-tokyo.ac.jp/en/ai1ec_event/13583/

[3] Second event: https://www.tc.u-tokyo.ac.jp/en/ai1ec_event/13803/

[4] Third event: https://www.tc.u-tokyo.ac.jp/en/ai1ec_event/14286/

1. Event Proceedings

Opening Remarks

The event began with opening remarks by Mr. Yoichi Iida (Ministry of Internal Affairs and Communications of Japan).

Mr. Iida noted that discussions on the EU-AI Act, particularly regarding the CoP, have entered a critical phase, and emphasized that the process is being advanced based on a multi-stakeholder approach, incorporating input from a wide range of actors including experts, industry, and civil society.

He went on to mention that there are growing concerns from various sectors that the current EU-AI Act may be overly burdensome. He shared an experience from a JICA training session where he lectured on AI governance to government officials from the Middle East, during which he was asked whether he thought the EU-AI Act constituted overregulation. While he refrained from giving a definitive answer, he expressed his personal view that the Act contains many excessive elements—an opinion that was met with a strong nod of agreement from the audience.

He also touched on last week's G7 discussions on AI governance, reporting that the advancement of the Hiroshima Process was a central topic and that the EU has shown an active stance toward it. At the same time, he raised concerns about the interoperability between the EU-AI Act and the Hiroshima Process, suggesting that there may be differences in interpretation and positions among the relevant stakeholders.

He concluded by emphasizing that excessive regulation is not desirable for the EU itself and expressed his hope that the discussions held at this event would be reflected in the ongoing debates within the EU.

Mr. Iida

Introduction to Panel Discussion: Overview of the public consultation and key points for Japanese companies to note

Next, Fumiko Kudo, Specially Appointed Associate Professor, Research Center on Ethical, Legal and Social Issues, the University of Osaka, and Professor Toshiya Jitsuzumi of Chuo University each provided explanations on the overview of the public consultation and key points for Japanese companies to note.

———

First, Specially Appointed Associate Professor Kudo provided an overview of the developments

surrounding the EU-AI Act to date, as well as an explanation of the newly released draft guidelines that are now open for public consultation.

At the beginning of her presentation, Specially Appointed Associate Professor Kudo explained that it is widely known—through media—that the EU AI Act is a hard law based on a risk-based approach, with four categories of risk defined according to the level of risk, that it has extraterritorial application and therefore also affects Japanese companies, and that violations can result in significant fines. She noted, however, that the main focus of this event is not on those four risk categories, but rather on a separate regulatory framework concerning "GPAI models." She explained that GPAI models refer to AI systems with broad applicability, such as generative AI and large language models (LLMs), which are connected to multiple applications or APIs. While these models are typically trained on large volumes of data, the definition is not strictly limited to such cases.

Specially Appointed Associate Professor Kudo also noted that AI models used for research and prototyping are, in principle, exempt from regulation, which serves as a reassuring point for researchers.

Next, she explained that obligations differ depending on whether GPAI models are considered to pose a "systemic risk." Systemic risks refer to concerns such as the lowering of barriers to the development of chemical or biological weapons, the emergence of AI systems beyond human control, and the spread of harmful discrimination or disinformation. If a model is not deemed to pose such risks, providers are required to fulfill obligations such as supplying technical documentation and establishing and implementing copyright compliance policies. However, if a model is assessed as posing systemic risks, more stringent requirements apply, including pre-deployment risk assessments and mandatory reporting of incidents to authorities. She also noted that cumulative computational power, measured in FLOP (floating-point operation), is used as a benchmark in this assessment—models **exceeding 10^{25} FLOP** are presumed to carry systemic risk. Additionally, she mentioned that EU authorities may designate models as high-risk on a case-by-case basis following consultations with experts.

Associate Professor Kudo reflected on her previous involvement in the development of the "CoP" for GPAI models, where she had provided input as an expert in earlier webinars. Initially, the CoP was scheduled to be finalized by May 2025 following multi stakeholder consultations, but as of the event date, it had not yet been completed. Meanwhile, she noted that a new and separate draft set of guidelines was unexpectedly released on April 22, 2025, with a public consultation open until May 22. Unlike the CoP, this new draft allows anyone to submit comments, and she encouraged active participation in the consultation process.

She also addressed the potential impact on Japanese companies, noting that entities involved in modifying or fine-tuning GPAI models may, under certain conditions, be regarded as providers—an issue with direct implications for Japanese businesses. Japanese experts had previously requested clarification on this point from the EU side, but it had been considered outside the scope of the CoP and remained unresolved. She pointed out that the new draft guidelines have now addressed this issue, reflecting Japan's request, and described this development as a positive outcome.

Associate Professor Kudo went on to explain that, based on the provisions of the EU-AI Act, the CoP has been working on templates for technical documentation as well as the development of safety and security

frameworks. She noted that the newly released draft guidelines include several key elements that Japanese companies should pay close attention to. These include a definition of "general-purpose", the boundaries between GPAI and specialized models, a proposed threshold of 10^{22} FLOP as a presumptive criterion for GPAI models (including those not posing systemic risk), and policies on how entities involved in modification or fine-tuning will be treated.

Associate Professor Kudo also offered her personal view on the background of these developments, noting that the EU is currently reviewing its AI regulations with the aim of strengthening its position in international competition. She pointed to several possible political factors influencing this trend, including the Draghi Report's[6] recommendations for regulatory simplification, pressure from the U.S. government, and calls for deregulation from French President Emmanuel Macron. At the same time, she emphasized that since the EU-AI Act has already come into force, making significant changes to its legal text is difficult. As a result, the focus is now on how flexibly the guidelines can be used to respond to these evolving challenges.

Finally, she pointed out that the draft guidelines include highly detailed technical descriptions, such as the formula for calculating FLOP and specific computation methods using NVIDIA chips (e.g., A100). She concluded by emphasizing that, for this consultation, input from business and engineering perspectives is more critical than that from legal or economic experts. She strongly encouraged participants to submit concrete suggestions for improvement.

———

Next, Professor Jitsuzumi provided additional explanations regarding the GPAI guideline draft, in line with the points raised by Associate Professor Kudo.

He began by noting that the provisions of the EU-AI Act regulating generative AI—particularly LLMs and GPAI—are scheduled to come into effect on August 2, 2025. In preparation for this implementation, discussions on the CoP had been underway since last year. However, he explained that the situation changed significantly following a speech by Vice President Vance in January, which triggered a shift in direction. As a result, the original schedule was significantly delayed in April, and a public consultation on the new draft guidelines was launched.

Regarding the upcoming schedule, he explained that the CoP was initially expected to be released in early May, but its publication was postponed. The current plan is to finalize both the CoP and the new guidelines by next month in order to meet the enforcement date of August 2. He noted that while the public consultation period has been extended, the preparation time before implementation has, in turn, been shortened.

He explained that the document is a non-binding supplement to the CoP and does not carry legal force. However, he pointed out that compliance with the guidelines can serve as a basis for asserting conformity with the obligations of the EU-AI Act. As such, for many companies and countries, adhering to the guidelines and the CoP may offer cost advantages.

---

[6] The Draghi report on EU competitiveness::
https://commission.europa.eu/topics/eu-competitiveness/draghi-report_en

He also noted that while the EU-AI Office had previously positioned the CoP as an implementation standard under the EU-AI Act—excluding interpretations of the Act itself—the latest guidelines, possibly influenced in part by negotiations with the United States, go a step further by including interpretive content regarding the Act. In particular, he highlighted that the guidelines address previously ambiguous issues, such as whether a fine-tuned model should be considered a new model depending on the extent of modification.

He further emphasized that the current public consultation is open to anyone and introduced the submission page for comments. He explained that the main points of discussion are practical in nature, including the definition of GPAI models, the extent of modification by downstream modifiers that would qualify them as new providers, the criteria for determining the time of market placement, and the acceptable scope of business models for open-source models.

In closing, Professor Jitsuzumi stated that, because these practical issues are difficult to address definitively from an academic perspective, input from those directly involved in providing services and building businesses is strongly encouraged.

Panel Discussion and Q&A

1. Discussion Points from Three Participating Companies

Before moving on to the panel discussion, three participating companies each presented key discussion points.

———

First, Mr. Kazuhiro Yoshinaga of NEC Corporation presented three key points of caution regarding the guidelines.

As an introduction, he explained the difference between the indicator "FLOP," as described in the current guidelines, and another indicator, "FLOPS." FLOP refers to the total number of floating-point operations and is comparable to an odometer in a car, measuring cumulative distance. In contrast, FLOPS refers to the number of floating-point operations per second and is akin to a speedometer, measuring instantaneous speed. The EU-AI Act uses the former—**FLOP**—as a measure of training volume for GPAI models. However, he pointed out that many explanatory articles in Japan and abroad tend to confuse the two, so caution is necessary. He also noted that earlier drafts of the EU-AI Act included the notation "FLOPs" with a lowercase "s," but it was removed in the final version to avoid confusion.

Building on this, Mr. Yoshinaga presented the first discussion point, focusing on the criteria for identifying GPAI models. The current draft guidelines set a training volume threshold of 10^{22} FLOP for determining whether a model qualifies as a GPAI model. He noted that most major AI models released since 2024 exceed this threshold, and therefore assessed the figure as reasonable, given that it does not significantly alter current eligibility judgments. For smaller models slightly below the 10^{22} FLOP mark, he explained that they typically do not reach one billion parameters and thus would not qualify as GPAI models under Recital (98) of the EU-AI Act.

The second discussion point concerned the criteria for determining whether an entity modifying a GPAI model qualifies as a provider. According to the draft guidelines, if the training volume required for

modifications—such as fine-tuning—exceeds one-third of the 10^{22} FLOP threshold, the entity may be regarded as the provider of a new GPAI model. Similarly, if the modification results in training that exceeds one-third of the systemic risk threshold (10^{24} FLOP), the entity may be classified as the provider of a new systemic-risk GPAI model. Mr. Yoshinaga noted that this framework helps clarify the scope of obligations and responsibilities and alleviates concerns that minor modifications might lead to full provider liability. However, he raised a concern regarding models like Meta's Llama3-70B, which already approach 80% of the systemic risk threshold even before modification. If such a model is fine-tuned and exceeds the threshold, the entity conducting the modification may bear responsibility for the entire model. Given that retraining existing large models is common practice in Japanese industry, he pointed out that this regulation could have a significant impact. To illustrate this concern, he presented a visual diagram showing how crossing the FLOP threshold alters the obligations of the modifier.

As the third discussion point, Mr. Yoshinaga addressed the criteria for "general-purpose (GP)" in the definition of GPAI models. He noted that even if a model exceeds the aforementioned training threshold, it may not be considered a GPAI model if the tasks it can perform are limited in scope. The draft guidelines explicitly state—with examples—that AI systems trained on specialized datasets or designed for specific tasks do not fall under the GPAI model category. Given that many Japanese AI companies focus on developing domain-specific or industry-specific AI systems, he emphasized that whether such specialized AI falls outside the GPAI definition will be a critical factor in determining the extent of the regulation's impact on Japan's industrial sector.

In conclusion, Mr. Yoshinaga summarized his three key points as follows:

First, the training volume threshold for GPAI models is reasonable and does not have a major impact. Second, special attention is needed when modifying models that fall just below the systemic risk threshold. Third, the criteria for determining whether specialized AI systems are excluded from the definition of GPAI models will be a crucial issue for Japan, where many companies are developing such domain-specific AI.

———

Next, Mr. Muneki Nemoto of NTT, Inc., who also serves as a specialist member of the CoP committee, presented his insights on the current draft guidelines.

Mr. Nemoto began by addressing the FLOP thresholds presented in the guidelines, expressing a positive assessment of the clear delineation between the systemic risk threshold (10^{25} FLOP), the GPAI model threshold (10^{22} FLOP), and the threshold for minor modifications to GPAI models (3×10^{21} FLOP). He also commended the guidelines for providing concrete interpretations of practical terms—such as "placing on the market"—noting that these clarifications are highly beneficial for real-world application.

A particularly noteworthy aspect, Mr. Nemoto highlighted, is the set of model examples—Models A through D—presented in the guidelines. Model A, which uses an NVIDIA A100 GPU, illustrates two different methods for calculating FLOP: the "hardware-based approach" and the "architecture-based approach." He praised the hardware-based approach in particular for its clarity, explaining that FLOP are calculated using the formula: training time (in seconds) × GPU performance × GPU utilization rate. He described this approach as highly intuitive and practical.

He provided a concrete calculation example, explaining that if a model is trained for 35,000 hours, the

resulting FLOP would be roughly twice the GPAI model threshold of 10^{22} FLOP—meaning it would exceed the defined criteria. If the training were conducted using 10 GPUs, this would equate to approximately 145 days, or just under five months. Based on this, he concluded that "when using 10 A100 GPUs, the GPAI model threshold would be reached in about 2.5 months."

On the other hand, Mr. Nemoto also expressed concerns about the FLOP-based thresholds. Specifically, he pointed out that a GPU's theoretical floating-point performance does not directly translate to actual training performance of AI models. For example, while the NVIDIA H100 offers roughly six times the computational performance of the A100, some reports indicate that its actual training performance is only about twice as high. As a result, he cautioned that as more powerful GPUs become available, the current FLOP thresholds may effectively become less stringent over time.

As an alternative, Mr. Nemoto proposed evaluating training volume based on the number of model parameters and training tokens. He suggested that this approach would be less susceptible to changes resulting from hardware advancements and could serve as a more stable and reliable metric over time.

He further recommended that the FLOP threshold should be periodically reviewed, noting that advances in algorithmic efficiency can significantly alter the amount of computation required for training. He also pointed out that the relationship between the volume of training and the societal risks posed by AI models remains insufficiently understood, highlighting this as an important issue for future discussion.

He also noted that the systemic risk threshold of 10^{25} FLOP is already established in the legislation and cannot be easily changed. Given this, he suggested that retaining the "architecture-based approach" formula—calculated as (number of parameters $\times$ number of training tokens $\times$ 6 $\approx$ FLOP)—could serve as a more realistic and practical alternative for assessing training volume.

———

Next, Mr. Kenta Oono of Preferred Networks, Inc. presented three key discussion points.

As the first point, Mr. Oono addressed the criterion that task versatility is a key factor in determining whether a specialized model is classified as a GPAI model. He noted that even models specifically designed for medical applications could potentially be considered GPAI models if they are capable of performing a wide range of tasks. He emphasized that this criterion could significantly affect the development of specialized AI systems.

As the second point, Mr. Oono raised a concern about the inclusion of computation used to generate synthetic datasets in the overall FLOP calculation for model evaluation. While both the working draft and the legislation state that the generation of synthetic data should be included in the FLOP count, he questioned whether this is truly appropriate. He expressed concern that if even synthetic data closely resembling non-synthetic data must account for the computational load of the model used to generate it, the use of synthetic data could become significantly restricted.

As his third point, Mr. Oono raised the issue of computational load during inference. He noted that techniques such as in-context learning in early-stage inference and the extended inference times of AI agents are increasingly contributing to improved performance. As a result, even with the same underlying model, general-purpose capabilities may expand depending on the computational cost of inference. He expressed concern that the current working draft does not clearly address how inference-side computation

should factor into model evaluation and stated that he will be closely watching how this issue is handled in future discussions.

In conclusion, Mr. Oono reiterated that his three main concerns are: the criteria for determining whether a model is considered a specialized model, the inclusion of computation used for generating synthetic datasets, and the role of inference-time computational load.

2. Panel Discussion

In the panel discussion, the three presenters who had shared their key points were joined by Specially Appointed Associate Professor Kudo and Professor Jitsuzumi, who had earlier explained the details of the public consultation. The discussion was moderated by Arisa Ema, Associate Professor at the University of Tokyo.

———

Specially Appointed Associate Professor Kudo began by expressing her appreciation for the insights shared by the three presenters. She recalled feeling confused when first reading the draft guidelines due to their technical content and noted that the preparatory meetings and the explanations provided by the speakers at this event had greatly helped deepen her understanding.

Specially Appointed Associate Professor Kudo then directed a question to Mr. Oono, referencing the three key issues he had raised for Japanese companies to consider: (1) the distinction between specialized models and GPAI models, (2) the treatment of synthetic datasets in FLOP calculations, and (3) the role of inference-time computational load. She asked whether it is correct to assume that, in the context of AI agent development, it would be advantageous for developers if inference-time computation is excluded from the evaluation of GPAI models. She also requested that he outline the possible positions or scenarios that could arise on this matter.

In response, Mr. Oono acknowledged that perspectives on the issue may vary depending on one's position. He pointed out that the higher the estimated computational volume, the more likely a model is to be classified as a systemic-risk GPAI model, which would in turn increase the regulatory obligations. Therefore, from the developer's standpoint, it may be preferable for inference-time computation to be excluded from regulatory assessments. He also noted that whether FLOP during inference should simply be added to the total is open to debate. Since the number of FLOP can vary significantly depending on the GPU used, he cautioned that evaluating models based on a straightforward addition of inference FLOP requires careful consideration.

In response, Specially Appointed Associate Professor Kudo thanked Mr. Oono for his insights and reiterated that the interests differ between companies developing LLMs or SLMs in-house and those using LLMs via AI agents. She emphasized that, since the current public consultation is open to everyone, it is desirable for a wide range of stakeholders to submit their perspectives, and concluded her remarks on that note.

———

Next, Professor Jitsuzumi commented that the earlier presentations revealed a shared emphasis among the speakers on the need for more detailed criteria. He noted that there may be two contrasting approaches:

one that favors greater specificity to narrow the scope of applicability, and another that deliberately maintains broad, ambiguous definitions—essentially a "least common denominator" approach—to make it easier for companies to argue that their services fall outside the scope of the GPAI regulations. He then posed a question to the panelists: from a business perspective, which approach is more preferable?

When Associate Professor Ema invited Mr. Nemoto to respond, he expressed appreciation for the current guidelines, noting that they provide a high level of specificity that is extremely helpful from a practical standpoint. However, he also raised concerns about the rapid obsolescence of such criteria due to technological advancements. In particular, he pointed out that changes in GPU types alone can significantly alter the thresholds, and therefore emphasized the need for criteria that are less susceptible to such shifts.

In response, Professor Jitsuzumi again suggested that while setting clear criteria based on current GPUs may make it easier to adapt to future technological developments, it could also become a constraint for Japanese companies down the line.

In response, Mr. Nemoto pointed out that under the current framework, technological advancements could effectively loosen the thresholds, leading to an expansion in the number of models subject to regulation. He argued that maintaining the criteria above a certain level—or reviewing them regularly—would be preferable from a business development standpoint. He further explained that even if the numerical threshold (10^{25} FLOP) remains unchanged, improvements in GPU performance would, in practice, lower the bar, thereby increasing the range of models that fall within the regulatory scope.

Finally, both parties agreed that, to prevent such situations, it would be advantageous from a business perspective to set the thresholds relatively high from the outset. This approach would allow the regulatory framework to remain effective and adaptable in the face of future technological developments.

Next, Associate Professor Ema noted that companies may adopt different strategic approaches, and invited the other panelists to share their views in response to Professor Jitsuzumi's earlier question.

Mr. Yoshinaga responded first, stating that the current gray area regarding the extent to which specialized AI models might be classified as GPAI models is quite broad, and that greater clarity would make it easier for companies to ensure compliance.

In response, Professor Jitsuzumi noted that under the EU-AI Act, regulatory criteria include not only computational thresholds but also model designations made by the EU-AI Office. He asked Mr. Yoshinaga to confirm whether it would be strategically advantageous for businesses to allow EU-AI Office to have such discretion.

Mr. Yoshinaga agreed with Professor Jitsuzumi's interpretation and added that another key issue is determining how narrowly an AI system can be scoped to a specific task while still being exempt. He reiterated that clarifying such gray areas would be beneficial for businesses.

Mr. Oono also shared his response, first noting that he largely agreed with Mr. Nemoto's perspective on the threshold criteria. He questioned, however, whether the current computational threshold of 10^{25} FLOP will remain appropriate in the future, emphasizing the need for periodic review. He pointed out that as

models become more efficient, smaller models can achieve accuracy comparable to much larger ones—for example, an 8B model trained on a well-curated dataset might match the performance of a past 100B model. Therefore, he stressed that the appropriateness of the thresholds should be re-evaluated in line with technological advancements.

Responding to Mr. Oono's comments, Professor Jitsuzumi summarized that since companies have varying strategies for future model upgrades, submitting opinions that address escape clauses could be a viable direction.

Mr. Oono agreed with Professor Jitsuzumi's summary, emphasizing that companies developing SLMs and those aiming for LLMs will likely have different responses.

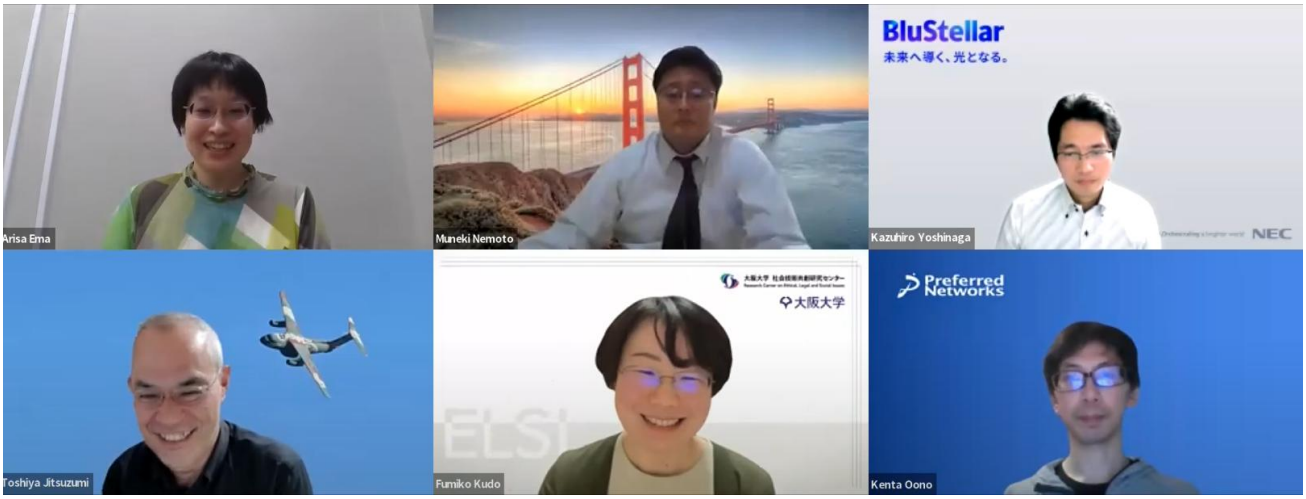
Finally, Professor Jitsuzumi stressed the importance of each company submitting their own public comments rather than relying on consensus with others. He called for active cooperation to ensure that the EU-AI Office understands the diversity of approaches existing in the industry.

The panelists and moderator then addressed several questions from attendees (for details of the discussion, refer to the Q&A section below).

In closing, Associate Professor Ema expressed her gratitude once again to the presenters for clearly articulating the key points within the limited time. She reiterated that the current public consultation is an open opportunity for anyone to submit their opinions. Furthermore, she emphasized that, since details may vary depending on each company's policies, proactive submission of comments is strongly encouraged.

She further noted that the main body of the CoP is expected to be published in June, and emphasized the need to continue monitoring the situation closely, as regular reviews of definitions and standards are planned going forward.

She also emphasized the great significance of gatherings like this one, which bring together diverse perspectives for collective reflection. Noting the ongoing development of AI legislation in Japan, she stressed the importance of continuing such opportunities for dialogue in the future. Finally, she expressed gratitude to the presenters and attendees for participating despite the short notice, and closed the event with anticipation for the next event.



Top row (from left to right): Associate Professor Ema, Mr. Nemoto and Mr. Yoshinaga

Bottom row (from left to right): Professor Jitsuzumi, Specially Appointed Associate Professor Kudo and Mr. Oono

2. Q&A

This section presents a selection of questions from attendees and the speakers' responses to them.

—— Judging general-purpose capabilities based solely on computational volume is inappropriate. According to the draft guidelines, the threshold of 10^{22} FLOP applies only to generative AI for image and text. Therefore, facial recognition AI, regardless of its computational load, would not be classified as a GPAI model. While the draft guidelines do not specify criteria for non-generative AI, if the current stance is that non-generative AI is generally excluded from GPAI models, this should be explicitly stated. Since all AI systems, including facial recognition, possess a certain degree of general-purpose capabilities, limiting the definition to generative AI alone does not alleviate concerns.

(Supplement) In the definition of GPAI under Article 3, item 63, the phrase "is capable of competently performing" might be determined by computational volume, but the requirements concerning "displays significant generality" and "performing a wide range of distinct tasks"—which are key elements—are not clearly articulated.

Specially Appointed Associate Professor Kudo:

I completely agree. As mentioned in the earlier presentation, since the EU-AI Act has already come into effect and its provisions are difficult to amend, targeting the current public consultation on this issue is almost meaningless. However, from a fundamental perspective, I strongly question how reasonably computational volume and generality relate to the systemic risks that the GPAI model aims to address—such as loss of human control, facilitation of weapon development, and spread of misinformation.

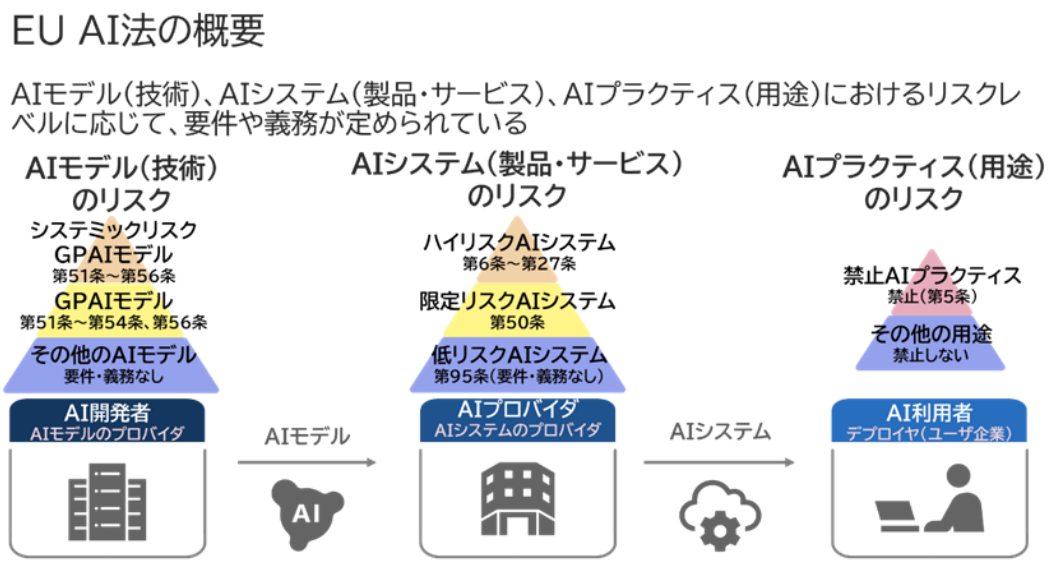
As widely reported, when California proposed a similar regulatory bill on frontier AI models (The Safe and Secure Innovation for Frontier Artificial Intelligence Models Act: SB1047), a similar debate arose regarding the rationality of setting thresholds based on computational volume in relation to the intended risks.

Constitutional challenges and other legal risks were also pointed out, which ultimately led to Governor Newsom's decision not to sign the bill, causing it to be shelved.
To reiterate, even though these provisions are embedded in the EU-AI Act, I strongly share the opinion that this approach is inappropriate.

Professor Jitsuzumi:
Indeed, while the current CoP and related discussions appear to focus solely on generative AI, the legislation itself does not explicitly limit GPAI models to generative AI. The definition provided refers merely to AI systems that are "trained with large amounts of data" and are "capable of serving multiple and varied purposes." Therefore, submitting public comments to clarify that the scope should be explicitly limited to generative AI can be a meaningful and effective contribution.

Mr. Nemoto:
The assumption that a high degree of general-purpose capabilities directly correlates with high risk—and that general-purpose capabilities can be assessed based on computational volume—lacks strong plausibility. Under such circumstances, the imposition of costly obligations may be a key reason behind the criticism directed at the CoP.

Mr. Yoshinaga:
(Regarding the supplement)
As Recital (99) lists only large generative AI models as examples of GPAI models, it is currently understood that GPAI models refer specifically to generative AI models. However, since AI models with general-purpose capabilities beyond generative models may emerge in the future, the term "GPAI model" is intentionally used with interpretive flexibility.
It should also be noted that under the EU-AI Act, obligations and requirements are imposed not only based on the risk of AI models, but also on the risks associated with the AI systems into which those models are integrated and the AI practices involved. Accordingly, even for AI models such as facial recognition systems that are not classified as generative AI (and thus not GPAI models), the associated products, services, and their use cases may still fall under regulatory requirements.



—— Is electricity consumption, etc. not subject to regulation?

Professor Jitsuzumi:

The Transparency section of the CoP requires disclosure of environmental performance, which is positioned more as a matter of information disclosure than regulation. Given this emphasis on transparency, it is highly likely that these standards will be referenced in public procurement processes across various countries in the future. If that becomes the case, instead of direct legal regulation, financial pressure may emerge as a form of regulation, as major users may choose not to adopt AI systems that fail to meet certain standards.

Mr. Yoshinaga:

As noted in Professor Jitsuzumi's response, the CoP requires information disclosure, which raises the possibility that it may be used as a reference standard in government procurement and similar contexts. In addition, a new law called the "Cloud and AI Development Act" is being considered separately from the EU-AI Act, with the aim of promoting the development of energy-efficient computing[7].
This legislation is expected to include the establishment of minimum energy-efficiency criteria [8].

—— Under the EU-AI Act, GPAI models targeting text are the primary focus, with algorithms like Transformers assumed as the baseline. However, image generation AI relies on different algorithms, such as diffusion models, which suggests that guidelines designed for GPAI models may not be directly

[7] AI Continent Action Plan - Q&A: https://digital-strategy.ec.europa.eu/en/faqs/ai-continent-action-plan-qa

[8] https://www.linkedin.com/posts/luca-bertuzzi-186729130_the-providers-of-cloud-services-for-ai-face-activity-7288525730085285888-vYJz/

applicable. This raises an important question: how should this be interpreted? It is essential to clarify whether image generation AI is considered free of systemic risk, or whether it is simply determined not to fall under the definition of a GPAI model.

Mr. Oono:
I agree. Given the significant impact of image generation on issues such as fake news, it should not be excluded from regulatory scope merely because it deals with images. This area appears to be underdeveloped; while there are rudimentary calculation methods—such as estimating how many tokens a single image represents when computing FLOP—comprehensive frameworks and in-depth discussions are still lacking. It is likely that further debate and regulatory development will emerge in the future.

Mr. Yoshinaga:
As stated in Recitals (99) and (105) of the EU-AI Act, generative AI models are subject to the regulation regardless of whether they generate text, images, audio, or video. Moreover, the current draft guidelines also include an example of FLOP calculation for image diffusion models in Annex A.1.

Mr. Oono:
I agree with Mr. Yoshinaga. While there is ongoing debate about whether computational load is truly an appropriate metric, its use likely reflects an underlying intention to define GPAI models in an architecture-agnostic manner.

—— Will criminal lists related to generative AI be created by the EU, or by an international body akin to an "AI police"?

Mr. Nemoto:
As for the development of criminal lists related to generative AI, no such initiative has been reported to date. However, since the CoP includes provisions prohibiting the use of pirated websites for training, it can be inferred that the handling of information from such sites is likely to be considered in future discussions.

—— Will these regulations change once technologies such as quantum computers become available?

Mr. Oono:
The authorities' position on this matter remains unclear at present. However, regardless of quantum computing specifically, advancements in technology—such as improvements in GPU performance, algorithms, and dataset quality—could change the implications of current regulations. Therefore, benchmarks like the 10^{25} FLOP threshold will likely need to be reviewed and updated on a regular basis.

—— When considering only training data, is there a risk that the data may become biased toward countries with large populations, such as China or India?

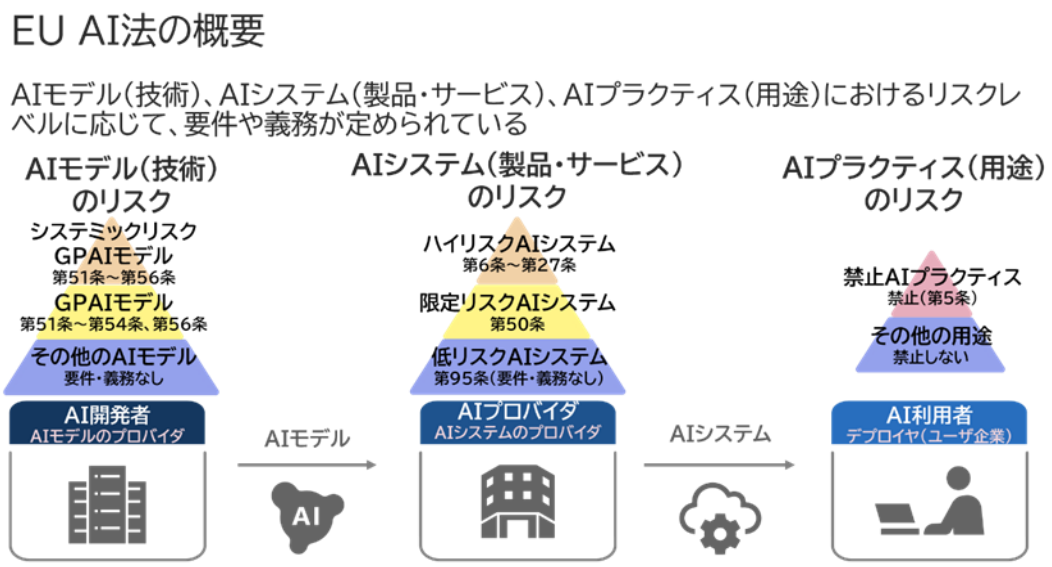
Mr. Yoshinaga:
From a volume standpoint, such bias is indeed possible, as countries with large populations may generate more data. However, in AI training, the quality of data is just as important as quantity. In fact, institutions like the National Institute of Information and Communications Technology (NICT) in Japan are actively working to extract high-quality Japanese-language data for training purposes[9].

Mr. Oono:
Population is likely to have a more direct impact on the number of users of a model, rather than on the number of organizations developing such models. However, considering factors such as the presence of companies (major corporations and tech startups) or research institutions with the capacity to develop large-scale models like GPAI models, the existence of national support systems, and the potential user base—especially for models trained in specific languages or cultural contexts—it is plausible that training efforts may become concentrated in economically stronger countries or regions.

—— Although still in the research and development stage, should vision-language-action models for robotics be assessed using the same computational thresholds?
If we assume that "the risk changes dramatically depending on the tools involved," is there a potential path where conditions are designed around the specific tools, environments, or scenes being handled?

Mr. Yoshinaga:
Since vision-language-action models enhance LLMs with image input and action (robot control) output capabilities, they are likely to fall under the definition of GPAI models. Regarding the varying risks depending on the tools involved, these may be addressed by imposing specific requirements or obligations on both the AI system (robotic product) and the AI practice (the intended use or application of the robotic product).

[9] https://www.soumu.go.jp/main_content/000997288.pdf



—— If computational load is to be accumulated, it would seem necessary to trace multi-stage histories. Are there any mechanisms being considered to monitor the supply chain in such a way?

Mr. Oono:

(With respect to the first half of the question)

In the current working draft, it is stated that "... the AI Office's preliminary approach is to allow providers to use reasonable estimates when precise information is impractical to obtain," indicating that when it is not feasible or practical to trace histories, providers are permitted to use estimated values as an alternative.

—— On page 5 of the guidelines, an example is given in which a "speech-only" application is still considered to fall under the definition of a GPAI model because it merely utilizes a subset of the capabilities of a broader model. This interpretation suggests that other applications, such as those for medical or academic research purposes, may also not be exempt if they are based on the partial use of general-purpose functions.

Mr. Yoshinaga:

Whether a model is considered an exception depends on whether it is seen as "a general-purpose model being used for a specific purpose" or as "a model that can only be used for a specific purpose," as described in the "Examples of models out of scope" on page 5. Given the ambiguity of this distinction, I raised the need for clarification during the topic presentation at this event.

Mr. Oono:

I agree with Mr. Yoshinaga. I believe further discussion is needed on the conditions under which a "medical LLM" would be considered as having general-purpose capabilities. For example, in addition to the model's inherent capabilities, factors such as whether the provider explicitly prohibits general-purpose use through

terms of service could be relevant considerations.

—— Regarding Meta's Llama 3-70B, which slightly falls below the systemic risk GPAI model threshold mentioned by Mr. Yoshinaga, is it possible that there are models with comparable accuracy but significantly lower FLOP counts?

Mr. Yoshinaga:
I agree. For example, when Gemma 2 27B was released last year, it was noted that its FLOP count and parameter size are only a fraction of those of Llama 3-70B, yet their performance is nearly equivalent.
Given this context, using FLOP as a sole indicator for systemic risk GPAI models remains somewhat questionable.

References

Related to this event

Public consultation submission form (Deadline: May 22, 2025, 12:00 PM CET)

【Video】Dealing with the Brussels Effect: How should Japanese companies prepare for the EU-AI Act? 4 (in Japanese)

Related to past events

【Event Report】Dealing with the Brussels Effect: How should Japanese companies prepare for the EU-AI Act? 3

【Video】Dealing with the Brussels Effect: How should Japanese companies prepare for the EU-AI Act? 3 (in Japanese)

【Event Report】Dealing with the Brussels Effect: How should Japanese companies prepare for the EU-AI Act? 2

【Video】Dealing with the Brussels Effect: How should Japanese companies prepare for the EU-AI Act? 2 (in Japanese)

【Event Report】Dealing with the Brussels Effect: How should Japanese companies prepare for the EU-AI Act?

【Video】Dealing with the Brussels Effect: How should Japanese companies prepare for the EU-AI Act? (in Japanese)

Report Preparation Assistance: Manaka Karino