Event Report

Building Trust through AI Governance and Transparency: Learning from the Hiroshima AI Process

Date	Thursday, 27 November 2025, 16:00–18:00 JST
Venue	Chiyoda-ku, Tokyo (Part I was also livestreamed on YouTube.)
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; Japan Deep Learning Association
Supported by	Cabinet Office; Ministry of Internal Affairs and Communications; The AI Safety Institute; The University of Osaka Research Center on Ethical, Legal and Social Issues

Introduction

The Hiroshima AI Process (HAIP) is an international framework for establishing AI rules, launched at the 2023 G7 Hiroshima Summit with the OECD Secretariat. In 2025, a "reporting framework" aimed at promoting transparency is scheduled to be fully implemented, and seven countries and 20 organizations have already published transparency reports. Companies such as Google, Microsoft, OpenAI, and Anthropic are participating, but notably, Japan has the largest number of participating companies, making it no exaggeration to say that Japan is taking the lead in international AI governance disclosure.

This event was held as a practical opportunity to learn about the significance and methodology of preparing such transparency reports. Using the handbook "Transparency Report Handbook for AI Governance" (Version 1.0)—developed primarily by the Ema Laboratory at the University of Tokyo—as the core teaching material, participants explored how to organize internal information, prepare reports, facilitate organizational consensus, and conduct effective governance meetings.

Event Report

The event consisted of two parts. In Part I, short presentations were delivered primarily by participating companies in HAIP, while in Part II, discussions on organizational governance were held among representatives from multiple companies, including the speakers from Part I. The following section introduces the content of Part I of the event.

Opening Remarks

At the outset, opening remarks were delivered by two speakers.

Mr. Takahiro Sumitomo (Counsellor, Secretariat for Science, Technology and Innovation Policy, Cabinet Office)

First, Mr. Sumitomo expressed his appreciation for the organization of the event and then outlined recent domestic and international developments surrounding AI. He noted that in Japan, "the Act on Promotion

of Research and Development, and Utilization of AI-related Technology (hereinafter, the "AI Act")"[1] was enacted in May 2025, followed by the establishment of the AI Strategy Headquarters in September of the same year, marking the full enforcement of the AI Act. Against this backdrop, he explained that discussions are currently underway toward the formulation of the Artificial Intelligence Basic Plan (hereinafter, the "AI Basic Plan") and related AI guidelines. He added that expert advisory panels have been convened as part of this process and expressed his gratitude to Associate Professor Arisa Ema (Associate Professor, Tokyo College, The University of Tokyo / Director, Japan Deep Learning Association), the organizer of this event, for her central role in these efforts. He also noted that the outline of the AI Basic Plan[2] was recently released and that discussions are now being deepened through the solicitation of public comments from a broad range of citizens.

He went on to emphasize that, going forward, it is important not only to decide on plans and guidelines but also to ensure their effective implementation, and that events such as this one play a significant role in that regard. He noted in particular that discussions within HAIP have entered a critical phase in which they must be translated into concrete initiatives, and he once again expressed his appreciation for the organization of this event.

He further noted that one of the four key basic principles set out in the AI Basic Plan is to take the lead in AI governance and enhance the trustworthiness of AI. Emphasizing that AI is a technology deployed across borders, he underscored the necessity of international governance that extends beyond Japan and highlighted Japan's position to lead discussions and framework-building in this area. He concluded that deliberations on how to concretely advance international AI governance going forward are therefore of critical importance.

Finally, he stated that the event would showcase a wide range of initiatives undertaken by various companies and noted that he was personally very much looking forward to the presentations. He also reaffirmed that the Cabinet Office is committed to pursuing trustworthy AI and working toward the realization of a country where AI is the easiest to develop and utilize in the world and called for the continued support and cooperation of all relevant stakeholders.

---

[1] The AI Act: https://www8.cao.go.jp/cstp/ai/ai_hou_gaiyou_en.pdf

[2] The AI Basic Plan: https://www8.cao.go.jp/cstp/ai/ai_plan/aiplan_eng_20260116.pdf

Mr. Sumitomo

Mr. Yoichi Iida (Advisor, Ministry of Internal Affairs and Communications)

Next, Mr. Iida took the stage, offering his congratulations and appreciation to Associate Professor Ema and the secretariat, and welcoming the strong turnout as a reflection of the high level of interest in AI governance. He stated that the Ministry of Internal Affairs and Communications has long been engaged in AI governance efforts, particularly on the international front, in addition to developing domestic guidelines, and that it continues to work in cooperation with the Cabinet Office—as noted by Mr. Sumitomo—toward realizing an environment in which Japan is the easiest country in the world for the development and use of AI.

He also explained that HAIP was agreed upon at the Hiroshima Summit in 2023 and, following a brief period of negotiations, was finalized by the end of that year. To ensure that the code of conduct agreed at that time would be effectively observed by companies, a reporting framework was established as a practical mechanism: by having companies report on their own initiatives, transparency is ensured, and users, markets, policymakers, and other stakeholders are able to access information related to safety. He noted that the aim has been to build trust through such voluntary efforts and to create a virtuous cycle in which trust leads to business value. He further reported that, domestically, 7 companies initially submitted transparency reports, and that more recently ABEJA, Inc. and Hitachi, Ltd. have joined the initiative, as well as that the increase in participating Japanese companies was highly commended at an OECD meeting held in Paris.

Looking ahead, he expressed the view that expanding the HAIP reporting framework as a global standard is a shared strategy not only for the G7 and the OECD but also for all relevant countries. In particular, he voiced his expectation that the HAIP Friends Group, scheduled to be held next March, will provide an opportunity to share this framework with a greater number of countries and to develop it into a mechanism in which companies around the world can participate. At the same time, noting that the reporting framework requires documentation in English and therefore places a considerable burden on Japanese companies, he expressed his gratitude for the substantial efforts and contributions made—led by the Ema Laboratory—in producing a Japanese-language “Handbook.” He emphasized that such a uniquely Japanese initiative has been highly regarded internationally.

Finally, he stated that the aim of the event is to share not only achievements but also challenges through

the initiatives of each company, and to continue making improvements so that the framework can take root as an established system. He also invited active issue-raising and feedback not only from the speakers but also from participants and concluded his opening remarks by expressing his expectations for the discussions to take place during the event.

Mr. Iida

Overview of Organizational AI Governance and the Handbook

AI Governance and Transparency

Next, in introducing the "Handbook,"[3] Associate Professor Ema addressed the significance of preparing and publishing transparency reports. She noted that while the act of producing a transparency report does not in itself immediately generate trust, it constitutes one of the necessary conditions for earning trust. She emphasized that trust-building requires organizations to carefully explain to society what kinds of AI they develop and how they govern them, and to gain the understanding of users and business partners and pointed out that transparency reports are positioned as a tool to serve this purpose.

She also addressed the question of why it is necessary to use the HAIP reporting framework when many companies already publish CSR reports, integrated reports, and IR materials. Regarding AI governance, she explained that companies are still in an exploratory phase as to what information should be disclosed and to what extent, and that this uncertainty is a shared challenge across organizations rather than one limited to specific companies. While the reporting framework sets out 39 items, she noted that it is not something that is completed once and for all; rather, it has the character of a "living document," in which the content to be shared evolves and expands in response to the emergence of new AI technologies and societal incidents or cases. She therefore emphasized that the ongoing collective effort to consider what should be shared with society is itself of critical importance.

She further stressed that the approach should not be reduced to a simplistic notion of "doing this ensures safety," but rather that companies and other stakeholders need to engage in ongoing dialogue to develop disclosure practices that are feasible for Japanese companies and to communicate information in ways that

[3] Transparency Report for AI Governance Handbook (Version 1.0): https://www.tc.u-tokyo.ac.jp/blog/wp-content/uploads/2026/01/HAIP_Handbook_EN.pdf

reflect what consumers in Japan and around the world actually value. She expressed the view that, through the accumulation of such continuous improvements, it becomes possible to advance the use of AI in a manner that goes beyond risk management and governance enhancement, achieving compatibility between innovation and trust as its foundation.

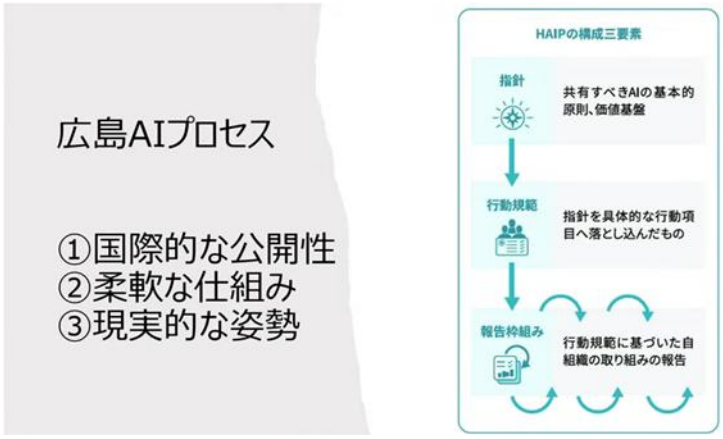
She then organized AI governance initiatives into a series of “phases” and explained the positioning of transparency reports within this framework. For companies considering the publication of transparency reports, she noted that they have already been engaged in efforts such as “Assessment”, “Vision Setting”, and “Organizational Structure” responsible for AI governance, and are now in a phase of considering how to communicate these efforts to society. Against the backdrop that many companies have already accumulated a certain level of knowledge and experience through trial and error in AI governance, she pointed out that there has been a lack of sufficiently concrete materials to support the next step—namely, guidance on “what to do next.” It was for this reason, she explained, that the Handbook was developed to support companies that have been working on AI governance in moving into the next phase of communicating their efforts to society.

Projection Slide (1)

HAIP

Next, she explained the positioning and key features of HAIP. She noted that HAIP is an international initiative developed within the frameworks of the OECD and the G7, and that a major distinguishing feature is the participation of many countries and regions—not limited to Japanese companies and the Japanese government—who share case studies and experiences.

She also emphasized that the reporting framework, consisting of 39 items, is a voluntary mechanism in which each company determines for itself what to disclose and how. Unlike hard law, where content and formats are strictly prescribed, the framework allows companies to indicate “not applicable” for irrelevant items and to disclose the status of future considerations for areas not yet addressed, reflecting a realistic and flexible approach. Accordingly, she explained that the essence of HAIP lies not in refraining from reporting due to incomplete systems, but in continuously demonstrating a commitment to future efforts and year-by-year growth and improvement. HAIP is therefore not a one-time submission, but a process that requires ongoing disclosure and iterative improvement.



Projection Slide (2)

In addition, she noted that, as components of HAIP, both guidelines and a code of conduct have been made public, and that the reporting framework—comprising 7 sections and 39 items—is positioned as a mechanism to support their implementation. She explained that the reporting content spans a wide range of perspectives, including risk identification, security, and organizational governance, as well as technical issues related to large language models (LLMs) and contributions to international and societal benefits. She emphasized that addressing these areas cannot be handled by a single department alone and requires cross-functional collaboration within organizations.

Projection Slide (3)

She also introduced that 24 companies have currently submitted reports under this framework, with Japanese companies being particularly active participants, followed by companies from the United States and other countries. She added that the number of participating companies is expected to increase further and expressed the view that being featured within such an international framework and introduced on the Ministry of Internal Affairs and Communications' website could itself serve as an incentive for companies.

Projection Slide (4)

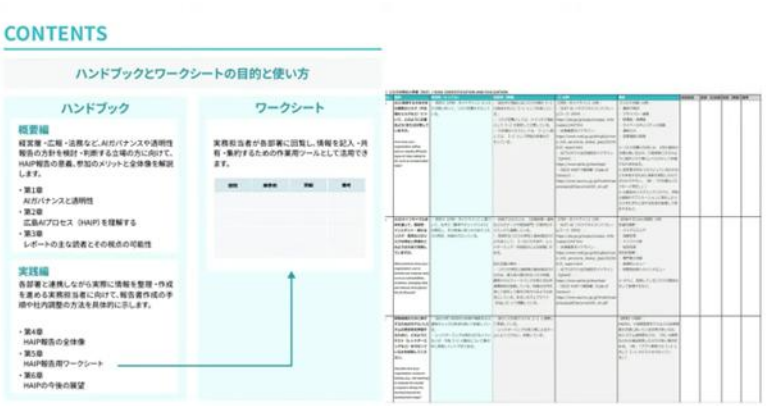
Contents of the Handbook and Worksheet

She then explained the structure and specific contents of the Handbook. She noted that the Handbook is composed of two parts—an "Overview Edition" and a "Practical Guide." The Overview Edition organizes key foundational issues, such as why AI governance needs to be considered, what HAIP is, and what benefits can be gained from participating in and reporting under the framework. This section is primarily intended for executives and management personnel.

By contrast, the Practical Guide is aimed at those who have been entrusted by executives and management with carrying out the actual work. She explained that it organizes practical considerations—such as how to structure internal teams, how to set agendas, and estimates of the required workload and time—based on interviews with companies that have submitted transparency reports. She emphasized that while the insights presented are general in nature, the section compiles content that is closely aligned with real-world practice.

She further introduced that, as a tool to support practical implementation, Excel-format worksheet[4] has also been prepared. To address the common question of "how to write," she explained that the worksheet provides templates in a question-and-answer format, informed by examples from companies that have already submitted reports. She added that the worksheet includes fields for entering the names of responsible personnel and response deadlines and expressed her hope that they would be actively used in practice by circulating and collecting them across relevant departments.

[4] Worksheet for Transparency Reporting on AI Governance: https://www.tc.u-tokyo.ac.jp/blog/wp-content/uploads/2025/12/HAIP_RF_Worksheet_en.xlsx (**When you click the link, the Excel file will be downloaded automatically.**)



Projection Slide (5)

Benefits of Reporting under HAIP

She then explained the benefits of reporting under HAIP. For companies seeking to demonstrate their commitment to AI governance on a global stage, she pointed out that using the HAIP reporting framework enables them to communicate their efforts in an internationally harmonized manner, rather than developing disclosures independently from scratch. She also noted that another advantage of such a framework is that it helps companies organize and visualize their own initiatives.

She further noted that, in the domestic Japanese context, the AI Act places importance on transparency and accountability, and while reporting under HAIP does not directly constitute compliance with the Act, it can nonetheless serve as a useful reference for companies in considering and preparing internally for what types of information may be subject to disclosure.

Projection Slide (6)

Conclusion

In closing, Associate Professor Ema concluded the session by expressing her hope that participants would join the discussions as members of the community, making active use of the worksheet.

Associate Professor Ema

Short Presentations by Participating Companies in HAIP

Next, it was announced that Fumiko Kudo, Specially Appointed Associate Professor (Research Center on Ethical, Legal and Social Issues, the University of Osaka), would serve as the moderator, and that 8 companies currently participating in HAIP would each deliver a one-minute presentation. It was explained that the aim of this session was to address the question of "what incentives motivate companies to participate in this reporting framework."

Three questions were posed to the participating companies: the first 3 companies were asked to speak briefly about the "Background and Motivations for Participating in HAIP"; the next 3 were asked about the "Benefits of Participating in HAIP"; and the remaining 2 were asked to discuss the "Approaches, Mechanisms, and Organizational Structures Devised or Being Implemented."

Specially Appointed Associate Professor Kudo

Background and Motivations for Participating in HAIP (1): Ms. Miho Naganuma (NEC Corporation)

As the first speaker, Ms. Naganuma of NEC Corporation (hereinafter "NEC") delivered a presentation. She began by noting that she had recently attended the OECD Digital Policy Committee (OECD DPC), chaired by Mr. Iida, where she presented views from an industry perspective. She also explained that NEC has participated in HAIP as a pilot since last summer and, together with NTT, Inc. (Mr. Muneki Nemoto), has been involved in preparations from the early stages of the framework.

She then cited two main motivations for participating in HAIP. The first was the belief that, given the

OECD's international influence and the strong involvement of Japanese government officials, discussions and institutional design within HAIP are highly likely to align with future Japanese policies. In an environment where the explanation of transparency and appropriateness are increasingly demanded, she explained that a key reason for participation was the perceived need for industry perspectives to be reflected from the policy and framework design stage.

The second motivation she identified was the importance of interoperability. For companies operating on a global scale, she noted that establishing internationally applicable common languages and frameworks—including those in English—is essential, and that NEC decided to participate in HAIP with the intention of contributing to the realization of such interoperability.

Ms. Naganuma

Background and Motivations for Participating in HAIP (2): Mr. Shinichi Kudo (SoftBank Corp.)

Next, Mr. Kudo of SoftBank Corp. (hereinafter "SoftBank") explained the company's initiatives, the background to its participation in HAIP, and its outlook. He introduced SoftBank's efforts to prioritize the social implementation and ethical use of AI under its guiding philosophy of "Information Revolution — Happiness for everyone," and outlined how, based on this philosophy, the company joined the HAIP partner community last year.

While generative AI has increasingly permeated many aspects of daily life and enhanced convenience, considerations of safety and transparency have not yet been sufficiently established. He explained that SoftBank decided to participate because it views the HAIP initiative as one that directly addresses the challenge of AI transparency and as an indispensable framework for the healthy and sustainable adoption of AI going forward.

He also stated that, through participation in HAIP, SoftBank aims to disseminate these initiatives not only externally but also internally, broadening understanding and practice around safety and thereby contributing to the wider adoption of better and more responsible AI.

Mr. Kudo

Background and Motivations for Participating in HAIP (3): Mr. Naohiro Furukawa (ABEJA,Inc.)

Next, Mr. Furukawa of ABEJA, Inc. (hereinafter "ABEJA") spoke about the motivations for participating in HAIP, as well as the insights gained and challenges encountered. He noted that a key trigger for participation was being invited by Associate Professor Ema.

As a positive outcome of the experience, he explained that after having been solely responsible for AI governance since joining the company in 2020, the process of preparing the transparency report allowed him to reassess and review ABEJA's initiatives. He noted that there were some questions for which it was difficult to determine how best to respond, and that working through these prompted him to recognize "aspects he had previously overlooked"—an experience he described as a significant takeaway. He also cited as a benefit the opportunity to systematically reorganize ABEJA's AI governance practices based on OECD principles, which deepened his own understanding as the person in charge.

On the other hand, as a challenge, he pointed out that some of the questions themselves were difficult to understand. In particular, he shared that concepts such as "tier" and "unreasonable risk" within the risk category were not always clear in terms of the intent behind the questions, making interpretation challenging. He remarked that, overall, navigating how to read and interpret the questions and making appropriate judgments required considerable effort.

Mr. Furukawa

Benefits of Participating in HAIP (1): Mr. Muneki Nemoto (NTT, Inc.)

Next, Mr. Nemoto of NTT, Inc. (hereinafter "NTT") explained the background to NTT's participation in HAIP and the effects it has generated. He noted that NTT established the position of Chief AI Officer last year and has positioned AI risk management as one of its core responsibilities. Based on the view that AI risk management is a domain of cooperation rather than competition, he explained that NTT places greater emphasis on understanding and aligning with the legal frameworks and societal expectations of its major business regions—such as Japan, the United States, and Europe—rather than pursuing a purely proprietary approach.

He also noted that HAIP is a framework in which Japan's international presence has been significantly enhanced, thanks in part to the efforts of the Japanese government, and that, within this momentum, he has been participating in discussions on the reporting framework since last year together with Ms. Naganuma of NEC. He explained that involvement in these international discussions enabled him to deepen his understanding of the reporting framework.

He further noted that, as the company has communicated its initiatives through opportunities such as event presentations, consultations from global companies—particularly those based in Japan—have increased, as have inquiries from institutional investors both domestically and internationally, and expressed the view that these developments are contributing to enhanced corporate value. He also explained that, internally, NTT is working to ensure comprehensive global AI governance across both geographic regions and the scale of the organization, and that strengthening major international collaborations such as those with the G7 and the OECD has helped build momentum and cohesion for advancing internal initiatives.

Mr. Nemoto

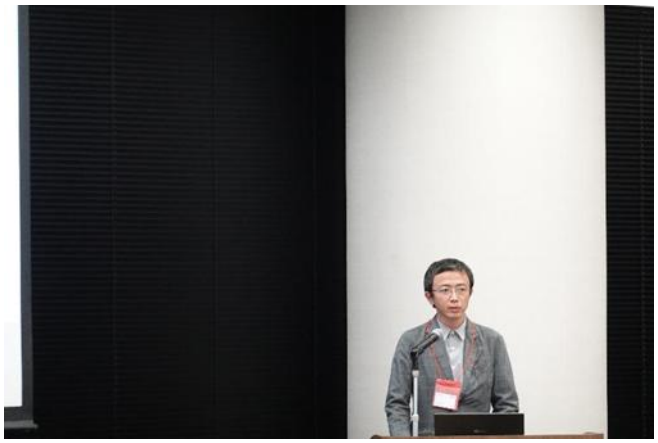
Benefits of Participating in HAIP (2): Mr. Akitsugu Ito (Rakuten Group, Inc.)

Next, Mr. Ito of Rakuten Group, Inc. (hereinafter "Rakuten") explained the company's AI governance initiatives and the significance of its participation in HAIP. He noted that Rakuten promotes a group-wide strategy under the banner of "AI-nization," advancing the use of AI across all areas of the business, and explained that the division to which he belongs is responsible in practice for establishing rules and managing risks in order to ensure that AI governance is appropriately implemented across the entire group.

He then identified two key benefits of participating in the HAIP reporting framework. The first was that, through the process of preparing the transparency report, Rakuten was able to foster a shared

understanding among both internal and external stakeholders of "what the company is doing with respect to AI." He noted that, in an environment where diverse fields and organizations are involved with AI in siloed ways, achieving cross-cutting understanding through the process of documentation was highly meaningful.

He cited as the second benefit the opportunity to comprehensively review AI governance and, in doing so, to rediscover that many areas could leverage or be adapted from existing IT governance measures. He noted that the process of taking stock of these measures and considering how to extend them to AI provided a valuable opportunity for further development.



Mr. Ito

Benefits of Participating in HAIP (3): Mr. Haruki Kojima (Microsoft Japan Co., Ltd.)

Next, Mr. Kojima of Microsoft Japan Co., Ltd. explained the benefits of participating in HAIP. He noted that the company has published a detailed transparency report of approximately 60 pages, which has also drawn considerable attention.

As benefits, he first noted, like other companies, that participation contributed to building a shared internal understanding and to taking stock of existing initiatives. In addition, he emphasized the importance of fostering a shared societal understanding that AI governance and AI utilization are not in conflict but rather should be pursued in tandem. He remarked that AI is a general-purpose technology, akin to electricity, and that for it to become widely embedded in society, it must be not only useful but also "trusted."

He also pointed out that, as AI technologies evolve from traditional AI to generative AI and further to agentic AI, there remains ongoing trial and error in determining how governance should be operationalized at the implementation level. In this context, he emphasized that the use of a shared framework and common language such as HAIP provides significant value by enabling internal and external stakeholders to engage in discussions on the same footing.

He further explained that, within an increasingly complex AI value chain, publishing transparency reports serves as an important means of communicating and bridging the company's initiatives to a wide range of stakeholders, including model developers and application developers.

Mr. Kojima

Approaches, Mechanisms, and Organizational Structures Devised or Being Implemented (1): Ms. Momoko Hosono (Fujitsu Limited.)

Next, Ms. Hosono of Fujitsu Limited. (hereinafter "Fujitsu") explained the approaches the company has devised and its internal structures in relation to the HAIP initiative. She noted that Fujitsu is celebrating its 90th anniversary this year and has been engaged in AI research and development for roughly half of that period, consistently pursuing "safe and trustworthy AI." She also explained that in 2019, Fujitsu established the "Fujitsu Group AI Commitment," centered on human-centricity, transparency, and privacy protection, and has been implementing it company-wide.

She stated that a major challenge in responding to HAIP was determining how to address the extensive set of questions spanning 7 sections and 39 items. As an approach, she described how she personally served as a hub, building a coordinating structure that brought together research divisions, departments responsible for ethics, quality, and security, as well as overseas offices. She also explained that, as priorities and perceptions of importance varied across departments, she took on the role—drawing on her position in external affairs—of organizing, translating, and aligning these perspectives to facilitate coordination.

She noted that while such coordination was by no means easy, the effort ultimately enabled Fujitsu to produce a practical, highly transparent report that clearly reflected the company's identity and strengths. She also emphasized that HAIP is not a one-off initiative; rather, Fujitsu is already working collectively across the organization toward the next update, continuously reflecting the outcomes in each department's operations and iteratively improving the content on an ongoing basis.

Ms. Hosono

Approaches, Mechanisms, and Organizational Structures Devised or Being Implemented (2): Mr. Koichi Takagi (KDDI CORPORATION)

Next, Mr. Takagi from KDDI CORPORATION (hereinafter "KDDI") introduced KDDI's organizational setup for responding to HAIP, as well as the points they had devised in the process. At KDDI, a project-based structure was adopted, in which a single coordinator was responsible for overseeing and consolidating the overall content of the report. He explained that he himself served in this coordinating role, managing the process and bringing together inputs from across the organization.

In addressing the initiative, one challenge was how to respond to the 39 items included in the framework. He explained that by replacing the phrase "your organization" in each item with their own company name and leveraging generative AI tools such as ChatGPT, KDDI was able to quickly produce draft responses. He also stated that AI-assisted drafting proved to be highly effective and significantly reduced the burden on the coordinator since HAIP requires companies to organize information that has already been publicly disclosed and restructure it in accordance with the questions. While additional effort was required to review and adjust the drafts through internal consultations, he emphasized that the ability to prepare an initial draft in a short period of time was a major advantage.

Finally, he offered practical advice to the participants, encouraging them to begin by first creating a simple draft as an initial step.

Mr. Takagi

Conclusion

In closing, Specially Appointed Associate Professor Kudo, the moderator expressed her appreciation for the presentations delivered by a total of 8 companies, noting that they provided highly practical advice. Then, the participants gave a big round of applause to all the speakers.

Associate Professor Ema also offered additional remarks, reflecting that the one-minute presentations by each of the 8 companies constituted a highly valuable and insightful session. She also informed the participants seeking more detailed information that they could consult the actual reports published on the OECD website[5] to review each company's specific initiatives.

She also noted that reading each company's transparency report can serve as valuable input for considering how one's own organization should engage with HAIP, as well as what information should be disclosed and in what manner. Furthermore, she encouraged the participants to recognize that, through efforts to ensure transparency, a shared international commitment is already emerging among many companies to view AI not merely as a risk, but as a technology to be appropriately leveraged while advancing innovation. The session concluded with renewed expressions of gratitude to both the participants and the speakers.

Closing Remarks

At the conclusion of the event, closing remarks were delivered by three speakers.

Mr. Mitsuhiro Hishida (Executive Director, the GPAI Tokyo Experts Support Center Secretariat)

Mr. Hishida explained the positioning of HAIP and GPAI (The Global Partnership on Artificial Intelligence)[6], as well as Japan's involvement in these initiatives. Drawing on his experience of having been engaged in G7-level AI discussions and the drafting of HAIP since his time at the Ministry of Internal Affairs and Communications, he expressed great satisfaction that, although the HAIP reporting framework is administratively demanding, participating companies have nonetheless evaluated it positively as "contributing to enhanced corporate value." He further emphasized that this initiative does not represent a one-sided imposition of burdens by the public sector but rather constitutes an internationally meaningful effort that contributes to strengthening Japan's global presence.

He also introduced the background and role of GPAI, which is closely linked to HAIP. GPAI was established against the backdrop of AI discussions within the G7 as a framework to support open and active discussions among experts from the public and private sectors and other stakeholders, centered on the theme of human-centric and trustworthy AI. He explained that GPAI currently operates through a three-center structure, with hubs in Paris (France), Montreal (Canada), and Tokyo, and noted that the GPAI Expert Support Center[7], established in July 2024, plays an important role as part of this framework.

He further emphasized that HAIP and GPAI share the common characteristic of being voluntary and flexible initiatives, in which experts participate on their own initiative and engage in open discussions on

[5] Submitted reports: https://transparency.oecd.ai/reports

[6] GPAI: https://www.oecd.org/en/about/programmes/global-partnership-on-artificial-intelligence.html

[7] GPAI Expert Support Center: https://www2.nict.go.jp/gpai-tokyo-esc/en/

issues for which clear answers have yet to be established, such as changes in work brought about by AI and the societal impacts of agentic AI. He also noted that, in the G7 Leaders' Statement, expert projects led by the OECD and GPAI are positioned as part of HAIP, and therefore supporting GPAI's activities also serves to advance HAIP.

Finally, he referred to a project on agentic AI that GPAI is currently prioritizing, noting that it plans to advance the preparation of reports based on scientific methodologies, including interviews with companies that have already introduced or are considering adopting such technologies. He added that there may be occasions in the future when cooperation from companies will be sought and called on stakeholders to work together to further strengthen and promote the HAIP reporting framework.

Mr. Hishida

In response, Associate Professor Ema commented that precisely because HAIP and GPAI are voluntary and flexible frameworks, they provide an important opportunity to discuss and communicate emerging and uncertain topics—such as agentic AI and LLMs—while sharing values and perspectives. Noting that the steady accumulation of such voluntary initiatives has contributed to enhancing Japan's international presence, she expressed her intention to continue these efforts without losing momentum going forward.

Mr. Satoshi Kurihara (Professor, Keio University / President, The Japanese Society for Artificial Intelligence)

Next, Mr. Kurihara raised key issues regarding the current state of global AI development and governance, as well as the unique role Japan can potentially play. At the outset, he noted that a paperback on AI ethics[8] is scheduled for publication in January 2026, highlighting that leading researchers in the AI field contribute discussions from a wide range of perspectives.

He then began by expressing a sense of unease about the current situation in which the AI and technology domains continue to be "dominated by strong players," while governance and rulemaking are being led by countries and regions that are relatively weaker in technological competitiveness, such as the EU and Japan. As Big Tech companies maintain their technological advantages and continue to compete, there is no

[8] Satoshi Kurihara (ed.), AI no Rinri Ningen to no Shinrai Kankei wo Tsukureruka (AI Ethics: Can We Build a Relationship of Trust with Humans?), Kadokawa Shinsho: https://www.kadokawa.co.jp/product/322410001045/

guarantee that a company that ultimately emerges as dominant will comply with established rules. On this premise, he pointed out the need to carefully consider what position should be taken in shaping governance and regulatory approaches.

He further argued that as AI acquires higher levels of autonomy and generality, it may transition from being merely a "tool" to an "entity to which humans delegate judgment." In such a phase, rather than "humans using AI", AI may begin to "transform humans themselves." He noted that the question of how far people should accept AI's judgments would then become a discussion about "trust" at an entirely different level from that of the past.

He suggested that while the acceptance of AI accompanied by transformations of human behavior and judgment may be difficult in Western societies, Japan may have unique potential to take on, debate, and implement such changes. The question of under what conditions people are willing to accept transformation brought about by AI is, he emphasized, a theme that Japan should address proactively. Rather than merely following discussions overseas, he underscored the importance of Japan taking the lead in articulating and disseminating its own perspectives.

Building on this point, he noted that if Japan were to independently design its own vision of AI and present models that ensure a certain standard of reliability, it could potentially hold a "keystone" role within the international division of responsibilities. In closing, he spoke of the significance of continuing to take on challenges even through voluntary initiatives, while keeping in mind a future in which Japan takes a leading role in rulemaking, and expressed his determination to move forward together with the participants.

Mr. Kurihara

In response, Associate Professor Ema noted that, even amid rapid change, the greatest value of a multistakeholder and voluntary community lies in the fact that it brings together people who choose to engage proactively, support one another, and move forward together. She also emphasized that the shared attitude of "building it together," rather than acting under instructions from others, is itself a significant achievement of this initiative.

Ms. Akiko Murakami (Executive Director, AI Safety Institute)

Next, Ms. Murakami discusses current trends in AI safety, the role of the AI Safety Institute (hereinafter

"AISI")[9], and future directions for collaboration. She expressed strong confidence in the fact that, in approximately year and a half since AISI's establishment, a forum has emerged in which many people have voluntarily gathered to engage in discussions centered on AI safety, emphasizing that this momentum must not be lost. She also likened the current situation to the early stages of the spread of the SDGs throughout society, noting that for companies and organizations utilizing AI, fulfilling responsibilities related to safety is increasingly becoming a taken-for-granted premise.

She further explained that AISI was established as a cross-ministerial initiative involving 12 government ministries and agencies, with the Information-technology Promotion Agency (IPA) serving as the secretariat, to support public–private collaboration on AI safety. Its primary mission, she noted, is to function as a hub for information related to AI safety. Beyond merely aggregating and disseminating information, AISI also aims to provide guidance on what companies and organizations should reference and the direction in which they should proceed.

As concrete outputs, she highlighted AISI's participation in the development of AI business guidelines, the formulation of guidance on red teaming and internal governance structures, and the open-source release of methodologies for evaluating AI models. She also pointed out the need to move discussions on AI safety—which have thus far remained relatively abstract— toward a more concrete and practical stage going forward.

She identified the importance of sector-specific discussions as a key factor going forward. Using the insurance industry as an example, she noted that AI-related risks unique to a particular sector should be shared and discussed as a non-competitive area, underscoring the value of considering safe AI utilization at the industry-wide level. To support such sector-based deliberations, AISI has established working groups beginning with the healthcare and robotics fields and has also launched cross-sectoral working groups to address common issues such as data quality and conformity assessment.

She also referred to the possibility of securing a certain level of government budget support for initiatives that entail costs, such as the development of testbeds. In addition, she noted that AISI aims to strengthen its organizational structure by expanding its personnel, and called for broad participation from diverse talent pools, including secondments and cross-appointments from companies and universities.

In closing, Ms. Murakami emphasized that the objective is not to put the brakes on AI, but rather to establish an environment in which it can be used with confidence, thereby providing the foundation that supports innovation. She noted that efforts to ensure AI safety ultimately enhance trust in companies and their services and stressed the importance of addressing rapidly evolving AI technologies not in isolation, but by sharing challenges and collaborating collectively.

[9] AISI: https://aisi.go.jp/about/

Ms. Murakami

In response, Associate Professor Ema identified “voluntary” and “trust” as the key themes that emerged from the event, reaffirming the importance of fostering both trustworthy AI and a trustworthy community. Amid significant shifts in technology, regulation, and society, she emphasized the necessity of connecting with trusted peers and continuing earnest yet engaging discussions. Expressing her expectation that each participant would begin acting within their respective organizations and roles, she brought the first session to a close.


Scenes from Part I (left) and Part II (right)

References

Related to this event

【Event Website】Building Trust through AI Governance and Transparency: Learning from the Hiroshima AI Process

【Video】Building Trust through AI Governance and Transparency: Learning from the Hiroshima AI Process (in Japanese)

Hiroshima AI Process

G7 reporting framework – Hiroshima AI Process (HAIP) international code of conduct for organizations developing advanced AI systems

The HAIP Reporting Framework: Feedback on a quiet revolution in AI transparency - Voluntary governance, stakeholder trust, and the future of responsible AI

Preliminary Insights and Recommendations for HAIP Reporting Framework - Based on HAIP Participant Organizations Interviews – (working paper)

Preliminary Insights and Recommendations for HAIP Reporting Framework - Based on HAIP Participant Organizations Interviews – (slides)

Transparency and Integrity in HAIP Reporting Framework - Insights and Recommendations for AISI Network – (slides)

Related to past events

【Event Website】What is the Hiroshima AI Process (HAIP) Reporting Framework? – Japan's cutting-edge global AI governance initiative

【Event Report】What is the Hiroshima AI Process (HAIP) Reporting Framework? – Japan's cutting-edge global AI governance initiative

【Video】What is the Hiroshima AI Process (HAIP) Reporting Framework? – Japan's cutting-edge global AI governance initiative (in Japanese)

Report Preparation Assistance: Manaka Karino