ACTIVITIES

AI Governance for "Trust from Society" in AI

Takuya Izumi
Director,
Information Policy Planning,
Commerce and Information
Policy Bureau, Ministry of
Economy, Trade and Industry

Tadashi Mima
Director,
Smart Infrastructure
Consulting Department (2nd),
Hitachi Consulting Co., Ltd.

Kohsuke Yanai
Department Manager,
Center for Technology Innovation
– Advanced Artificial Intelligence,
Research & Development Group,
Hitachi, Ltd.

Action is starting to happen on putting AI principles into practice. Japan's METI published a set of AI Governance Guidelines in July 2021 to support practical implementation of the AI principles needed to encourage the wider adoption of the technology. Hitachi meanwhile, has formulated its own principles for the ethical use of AI and has started work on their deployment. At the Hitachi Social Innovation Forum 2021 JAPAN held in October 2021, two key staff from Hitachi were joined by METI's Takuya Izumi, one of the leaders involved in preparing the ministry's AI Governance Guidelines, to share their views on how the many stakeholders from government, business, and elsewhere should coordinate their efforts to put AI ethics into practice.

Presentation 1
Managing Risk and Maximizing Positive Impacts
METI AI Governance Guidelines   (Takuya Izumi)

Artificial intelligence (AI) is becoming increasingly important as a source of innovation and it was against this backdrop that the Ministry of Economy, Trade and Industry (METI) published its AI Governance Guidelines* on July 9, 2021. I would like to start by talking about what led up to the publication of these guidelines.

---

* Ministry of Economy, Trade and Industry, "Governance Guidelines for Implementation of AI Principles Ver. 1.0"
https://www.meti.go.jp/shingikai/mono_info_service/ai_shakai_jisso/pdf/20210709_9.pdf

AI offers considerable benefits for society. Image recognition AI, for example, is used in the development of autonomous driving and in the interpretation of X-rays. On the other hand, if discrimination or other forms of bias are inherent in the training data used to build AI systems, those biases may be exposed by the AI system. The history of work on AI governance is one of seeking to enjoy the benefits that AI brings while also dealing with its negative consequences.

To advance this work, the Japanese Cabinet Office published a set of high-level guidelines entitled "Principles of Human-centric AI Society" in March 2019. This document defined seven such principles, including privacy, security, fairness, accountability, and transparency. Debate within the government on AI governance for facilitating work that follows these high-level guidelines began in earnest in June 2020. METI released a document entitled "AI

Governance in Japan" in January 2021 and called for comment. The METI considered the feedback received and published AI Governance Guidelines in July.

In this work, METI defines "AI governance" as "design and operation of technological, organizational, and social systems by stakeholders for the purpose of managing risks posed by the use of AI at levels acceptable to stakeholders and maximizing their positive impact." I would like to explain this further in the context of social systems.

Looking at **Figure 1**, the values sought after by society and embodied in the AI principles are listed under the "What" heading. The current challenges are shown under "How." In practice, there are a variety of ways in which we could go about fulfilling the social principles of AI. One possibility would be to introduce industry-wide (horizontal) rules that focus on the distinctive functions of AI. We could also establish technical standards. Yet another option is to introduce rules specific to particular sectors such as automobiles or healthcare. The challenge for AI governance is how to combine these in a way that maximizes the positive impacts of AI while minimizing the negative ones. Having considered industry-wide rules, METI has concluded that it is better to issue guidelines that are not legally binding. This led to the formulation of our AI Governance Guidelines.

The METI AI Governance Guidelines provide guidance on how to implement AI governance at the corporate level. The guidelines also draw on METI's agile governance framework. This framework applies to the establishment of governance practices in fields such as AI where the technology is still evolving and the nature of public acceptance is still changing.

Figure 2 shows AI governance as a nested loop. It begins with a condition and risk analysis. In the case of AI, what matters is to understand the technology, public acceptance, and one's own circumstances. This analysis is then used as a basis for setting goals. This involves determining which values are most important to the companies concerned. After this, the next step is to establish AI management systems for achieving the goals. To achieve these goals in the case of AI, it is important to conduct an as-is/to-be analysis to determine how the current situation differs from the goals. Also important are to improve literacy and to encourage businesses to work together. The implementation shown at the bottom of the loop relates to monitoring of the AI management system and individual AI systems. Appropriate external disclosure is also important. The evaluation in the inner loop looks at whether the AI management system is functioning as it should, and, if not, how it should be improved. Furthermore, if the business environment changes significantly, it will become necessary to perform a new condition and risk analysis and to repeat the process from the goal-setting step. The AI Governance Guidelines are designed to allow companies to engage in continuous improvement through dialogue with stakeholders. As one of the people involved in their formulation, I look forward to seeing these guidelines put into practice, even if only gradually.

Figure 1 │ Overview of AI Governance

The societal aspects of AI governance can be understood in terms of a hierarchy made up of: (1) Goals, (2) Horizontal and intermediate rules, (3) Rules focused on specific targets, and (4) Monitoring and enforcement.
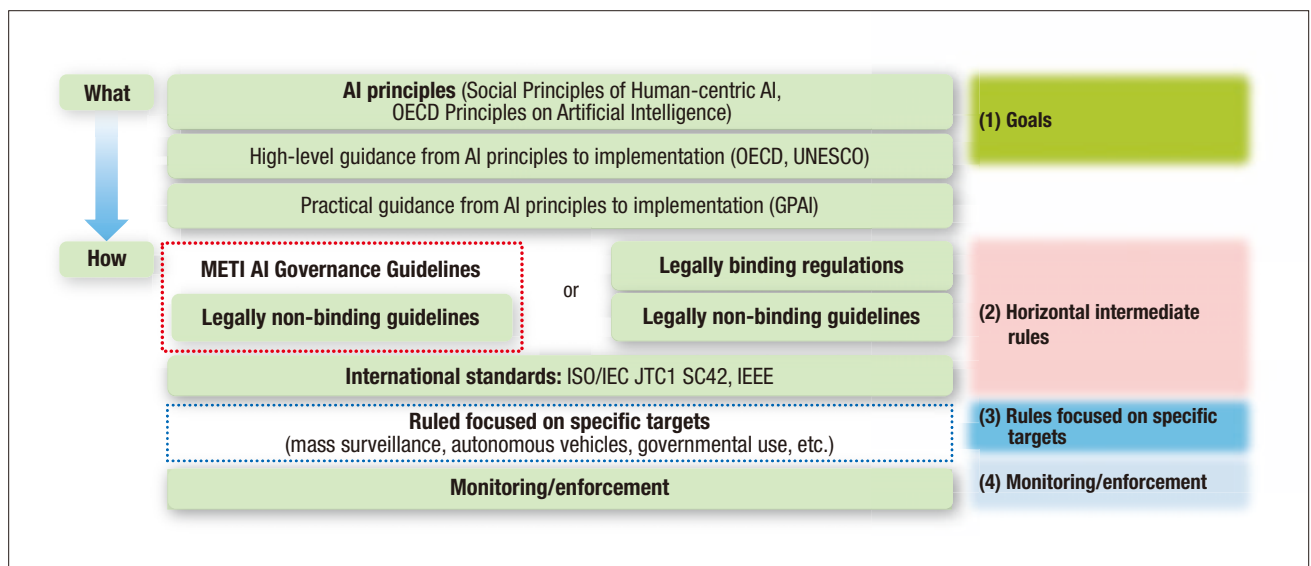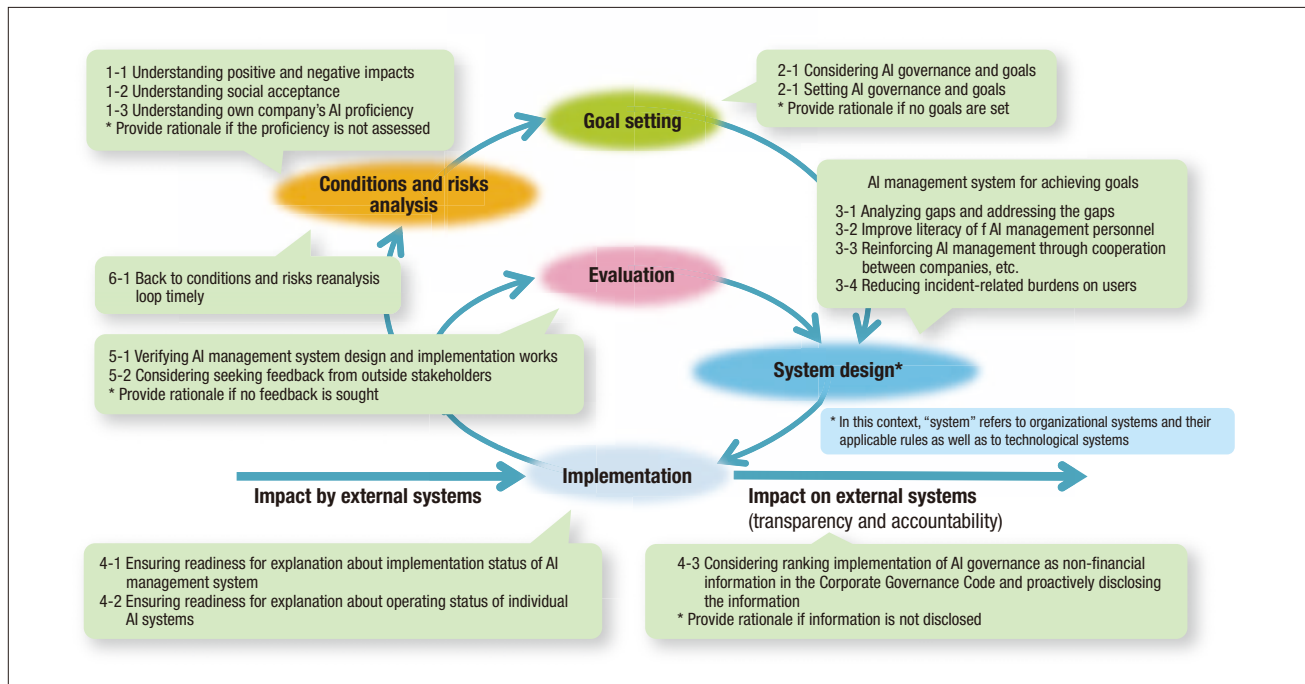
Figure 2│Overview of METI AI Governance Guidelines

The METI AI Governance Guidelines provide guidance on how to implement AI governance at the corporate level.

Presentation 2

Aiming for AI Ethics by Design
What Hitachi is Doing on AI Governance   (Tadashi Mima)

Hitachi has also started to take action on AI governance. Recognizing that AI is such a rapidly moving field, the question of how to establish governance practices that are both flexible and fast is of particular importance. Today I would like to talk about what Hitachi is doing in this area in the context of METI's agile governance framework.

First of all, in terms of the goals in **Figure 2**'s nested loop, Hitachi has established its own principles for the ethical use of AI that draw on what we learned from working through the analysis of AI situation and risk. One feature of these principles is that the setting of goals should be informed by an understanding of the AI lifecycle. Drawing on our considerable experience of working in areas such as social infrastructure that require long-term operation, Hitachi has set criteria for action that take account not only of those items to be addressed that apply generally, but also to the separate planning, societal implementation, and maintenance and management phases.

The next step after this is to establish the organization, people, and practices needed to operate the system design that in turn provides the governance to achieve the goals that have been set (see **Figure 3**).

The Lumada Data Science Laboratory (LDSL) is an organization within Hitachi that employs data scientists. In terms of specific measures, Hitachi has brought together a range of experts to form a specialist AI ethics team within the LDSL and has established mechanisms for coordinating with other groups across

the company, including our privacy protection advisory committee. An AI Ethics Advisory Board made up of outside experts has been established to accumulate knowledge as well as to ensure objectivity. In addition to working in partnership with this advisory board, the AI ethics team functions as a center of excellence (CoE) with a role that encompasses information sharing, education, and training.

Operationally, the work involves the assessment of AI ethical risks, the conducting of evaluations based on checklists and by the AI ethics team and advisory board, and the implementation of actions and other measures identified on this basis. Based on the concept of "AI ethics by design," Hitachi is seeking, wherever possible, to achieve the goal of reviewing and acting on its principles guiding the ethical use of AI from the planning and design stage or from the proof-of-concept (PoC) stage.

Finally, the evaluation in the inner loop involves periodically taking time out to review the system (at intervals of about six months) and to make improvements besides day-to-day operations. To achieve this, the AI ethics team as a CoE engages in continuous learning on topics such as policies and technologies that relate to AI and ethics.

Presentation 3

AI Ethics for Hitachi's Social Innovation Business
Research, Development, and Deployment of AI by Hitachi
(Kohsuke Yanai)

Hitachi's Social Innovation Business deals with social infrastructure and services with a strong public interest component, extending from health, medicine, and care to transportation and finance. Accordingly, being aware of just how much its activities can impact society, Hitachi has an obligation to take ethical factors

Figure 3 Organizations, People, and Practices in System Design

Consolidate knowledge in a Center of Excellence (CoE), including the establishment of a dedicated organization, training, and working with external partners.
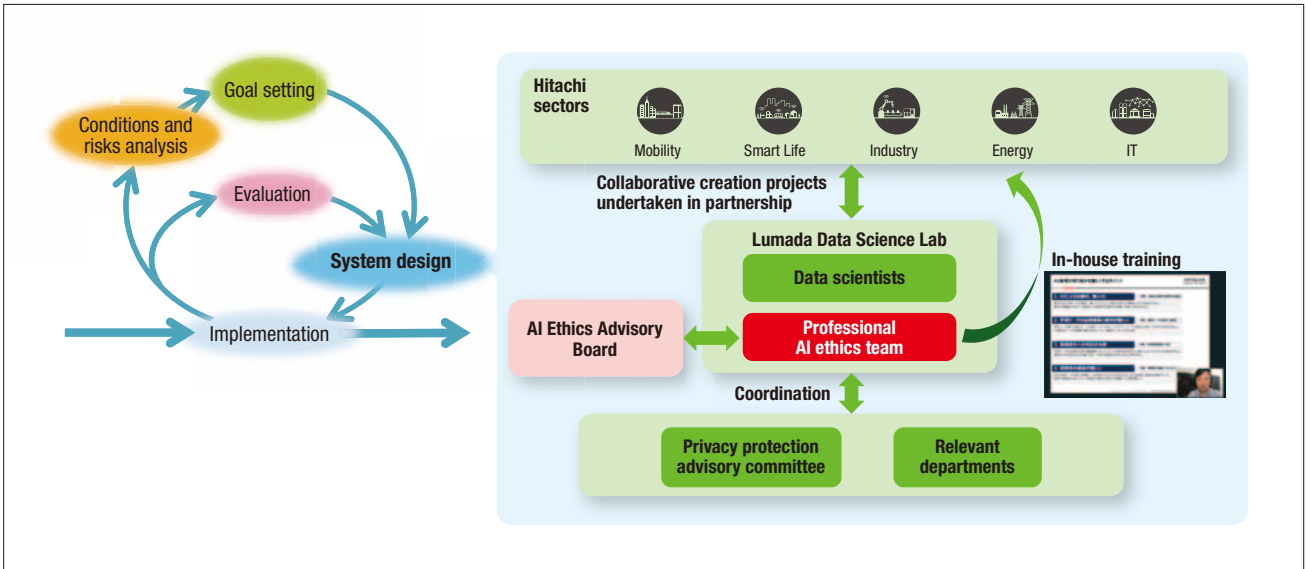
into account in the development and utilization of AI in its businesses. Today I would like to tell you a little more about the practical steps Hitachi has been taking in this area.

Figure 4 shows an example of how Hitachi goes about technology governance. Further details can be found in the article "Principles and Practice for the Ethics Use of AI in Hitachi's Social Innovation Business" (page 16 of this issue). Among the key factors emphasized at Hitachi are: (1) Establishment of in-house governance practices, (2) Training of staff to develop and use AI correctly, (3) Dialogue and consensus-building with stakeholders, and (4) Research and development of technologies that underpin AI ethics. To provide a common set of guidelines and terminology that can be used in these activities, Hitachi has also formulated its own

principles guiding the ethical use of AI in its Social Innovation Business.

Without a deep understanding of their technical characteristics and the changes they impose on society, new technologies such as machine learning and other forms of AI are at risk of being misused. Accordingly, a deep understanding of AI itself is a prerequisite for putting AI ethics into practice and this is something that needs to be addressed systematically and over the long term.

Hitachi has experience engaging in dialogue and consensus-building with stakeholders. This has included the establishment of the Hitachi-UTokyo Laboratory at the University of Tokyo and our work in partnership with Chulalongkorn University in Thailand. Recognizing the importance of talking to and coordinating with

Figure 4 How Hitachi Wants AI to Function in Society and How It Goes about Achieving This

Hitachi wants to create a world in which people trust AI and where it has the consent of the public who feel confident about its use to take on a variety of tasks.
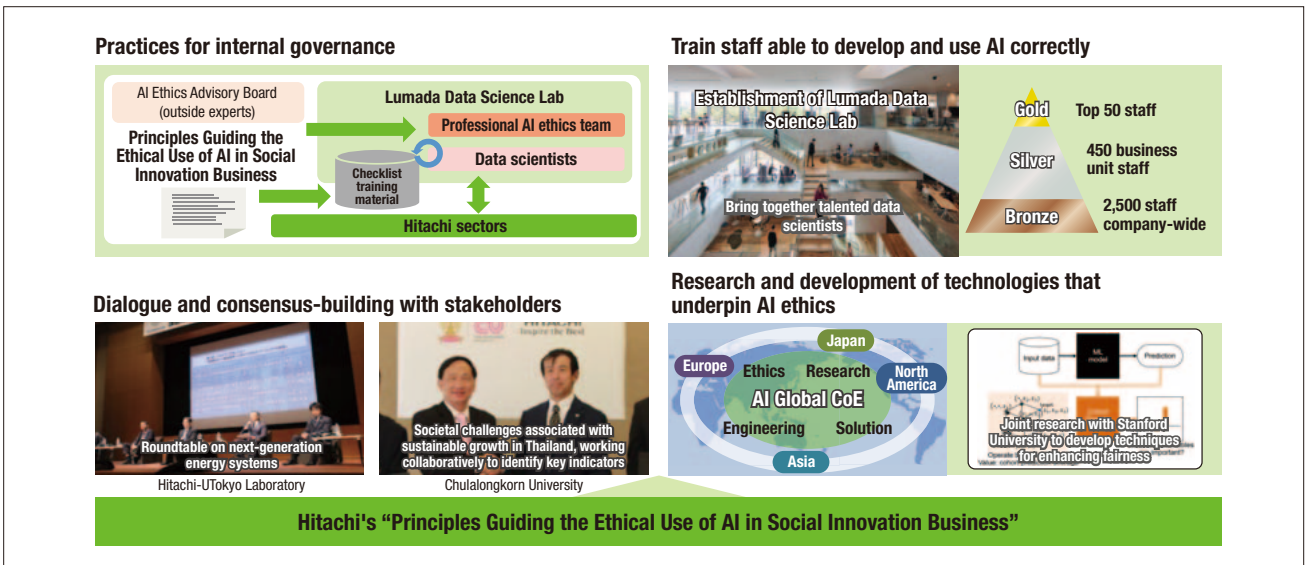
multiple stakeholders in the research and practical deployment of AI, Hitachi is stepping up its efforts in this area.

Hitachi has also established an AI Global CoE for the research and development of AI at facilities located across Japan, Asia, North America, and Europe. This involves going about research and development in an inclusive manner that takes diverse viewpoints into account. AI ethics covers a range of aspects, including safety, fairness, transparency, security, and privacy, and Hitachi's research work is likewise broad in scope. In one example, Hitachi is working in partnership with Stanford University on world-leading research into techniques for enhancing fairness using explainable AI (XAI). Along with research into AI itself, other work is looking at data governance and at trust governance. One such example involves the development of a system for optimizing the deployment of emergency services that use XAI. Services with a strong public interest component require high levels of fairness and explainability. Accordingly, it is by utilizing techniques designed to meet this need that successful practical implementation has been achieved.

Discussions
Coordinating AI Governance with Multiple Stakeholders

——Please give your views on the importance of the AI Governance Guidelines and how you would like to see them used in the private sector.

Izumi: Two things in particular are significant about the AI Governance Guidelines. The first is the support of the companies involved. A set of 14 action targets have been laid out that companies should be undertaking to facilitate the practical application of the AI principles needed to encourage widespread adoption of AI. The supporting material provided to accompany these targets includes virtual examples of each of these actions being put into practice and also examples of gap analyses of AI principles. The virtual action examples provide specific information, as do the gap analyses, and I believe that these will be a useful resource for

companies in their implementation of AI governance. The second is to underpin the actions that companies take at their own initiative by making extensive reference to the AI principles in situations such as the contracts entered into by companies in their development and operation of AI systems, and through the building of mutual understanding between the stakeholders involved in putting the principles into practice. Many different participants are involved in the development and operation of AI systems before they reach the stage of being used by their ultimate end users. Unless all of the companies in this value chain take the same approach to putting AI governance into practice, the negative impacts of AI might make themselves known. Our hope is that the AI Governance Guidelines will provide a common vocabulary through which the many participants can align their viewpoints. Given your position at the forefront of both the AI governance debate and its implementation, I hope that Hitachi can serve as a leader and as a hub bringing different companies together as they put AI governance into practice, and that you can work to ensure that society as a whole is able to enjoy the benefits of AI while also minimizing its negative impacts.

Mima: As you say, AI governance is not something that one company can achieve on its own. At Hitachi, we recognize our corporate clients as being the users of AI and understand the importance of communicating with them to enable the principles guiding the ethical use of AI to be followed in practice across all businesses that use the technology.

As in the example set by METI, broad applicability plays a large part in how we approach AI governance, and achieving a commonality of approach across the industry is vital. Hitachi is looking for ways of working with other companies who are addressing the issue of AI ethics, engaging with them either one-on-one or through industry organizations.

——It has been noted that agile governance is well suited to the technology of AI. Can you tell us a little more about why this framework is needed?

Izumi: Cyber physical systems (CPSs) like AI represent a high-level fusion of the cyber and physical worlds. Risk will likely prove

difficult to control in societies built on such systems, which will also be complex, fast-changing, and unpredictable. Moreover, the goals of governance will likewise continue to evolve in response to this societal change. These ever-changing circumstances and goals will in turn call for constant revision of how best to solve the challenges facing this society. This is why I do not believe that rigid governance models based on predetermined goals and methods are the right way to go. Agile governance provides a framework for the ongoing revision of what constitutes an optimal solution. In particular, I believe that the process of identifying how the current situation falls short of the goals, and taking action to resolve these gaps, needs to be seen as an integral part of the guidelines. Moreover, these gap analyses need to include people who are not directly involved in the development and operation of AI systems. What we do not need, I believe, are measures that block the development and delivery of AI purely because of the existence of these gaps. Gap analyses are no more than opportunities for improvement, and as such what matters is to work through the full loop. Also of importance is to provide users with adequate information about the potential for gaps and what to do about them.

Yanai: In practice, I agree with this even if only looking at it in terms of the research and deployment of AI. While the rapid changes in society and the rapid advance of technology are at the heart of it, there is no point to the systems we are creating unless they are genuinely needed by society and trusted by the public. While Hitachi, in operating its Social Innovation Business, needs to take responsibility for technology governance, it also needs to engage in dialogue with the different stakeholders involved, including government agencies, academia, communities, and the general public, and to build trust between these stakeholders as we go about our work. This is another reason why I believe it is important to work through the agile governance loop.

——Specifically, in what ways should the many stakeholders coordinate with one another in implementing AI governance and putting it into practice?

Izumi: If stakeholders are to work together, they need a common vocabulary. It was to provide this that we created the AI Governance Guidelines, and it is important that we use them to foster mutual understanding at the societal level. Coordination between stakeholders can take many forms and there is no magical one-size-fits-all solution. I believe we need to work on making progress on all of them.

Yanai: I see this as an area where we are still finding our way. We need to be continually revising governance in a flexible, ongoing, and timely manner. In doing so, collating the diverse views of stakeholders is vital. In its work toward realizing Society 5.0, Hitachi has been conscious of the need for dialogue and consensus-building with stakeholders. Drawing on this experience, we want to be even more attentive to bringing multiple stakeholders onboard in our work on AI. As a first step for ourselves as a company, I believe it is important that we go further than we have in the past in terms of being accountable to stakeholders.

Izumi: Something that is often said of Japanese companies is that they rarely take the initiative in setting rules. While the companies that make the rules often bear the brunt of criticism, it is important that we act resolutely. I would like to see Japanese companies actively involving themselves in these activities. Those who take the initiative are also among the first to identify the risks. Likewise, when the companies are subject to criticism, I would like to see them getting on the front foot and responding proactively.

——Finally, can you say a few more words about what you are looking for from Hitachi?

Izumi: As both a leader and a hub for the practical implementation of AI governance, I hope that Hitachi will help show the way for all stakeholders to work together and reach a common understanding, including business partners and end users. We face challenges that cannot be overcome by one company on its own. I also hope to see you actively engaged in opinion sharing and other such activities with corporate customers, partners, and people working in other sectors with the aim of involving yourself in the rule-making process.

Yanai: That's right. We intend to continue working hard to play a leadership role in the adoption of AI governance, taking input also from the outside experts who make up Hitachi's AI Ethics Advisory Board. With regard to rule-making, we also hope to contribute to the formulation of de facto standards that achieve public acceptance, including by making our views known to the relevant communities and institutions and through participating in demonstration projects that involve stakeholders.

——Thank you for your time today.