

Study Group 'AI governance and its Evaluation'
Report on the Session #9

1. Introduction

The Japan Deep Learning Association establishes study groups as a forum for deepening knowledge and discussing domestic and international policy trends related to artificial intelligence (hereafter AI) and Deep Learning (hereafter DL). This study group, "AI Governance and its Evaluation," defines "governance" as a system of management and evaluation by various actors, and launched a study group in July 2020 to investigate what forms of governance are possible and conduct a year-long study to help build trustworthy AI systems.

In the 9th session (February 18, 2021), Mr. Hiroki Habuka and Mr. Hirofumi Sugano, Information Economy Division, Commerce and Information Policy Bureau, Ministry of Economy, Trade and Industry (METI), presented topics under the views of AI Governance in Japan and METI studies in the area of Governance Innovation.

This report is a reconstruction of the topical presentations and the discussions of the study group participants.

2. Guide to Designing and Implementing Agile Governance

In the first part of the session, Mr. Habuka introduced an overview of the draft report "Governance Innovation ver.2: A Guide to Designing and Implementing Agile Governance¹" released on February 19, 2021² by the "Study Group on New Governance Models for Society 5.0" of the METI.

Overview of Society 5.0 and its issues

Society 5.0³, which is the premise of agile governance, is a concept of the future society proposed by the Japanese government since 2016, and it refers to a human-centered society that balances economic advancement with the resolution of social problems by a system that highly integrates Cyber-Physical System (CPS)⁴. In Society 5.0, the analysis results obtained by AI will directly affect the systems in the physical space, and

¹ <https://www.meti.go.jp/press/2020/02/20210219003/20210219003-1.pdf> (in Japanese)
https://www.meti.go.jp/english/press/2021/pdf/0219_004a.pdf (in English)

² <https://www.meti.go.jp/press/2020/02/20210219003/20210219003.html> (in Japanese)
https://www.meti.go.jp/english/press/2021/0219_004.html (in English)

³ https://www8.cao.go.jp/cstp/society5_0/ (in Japanese)
https://www8.cao.go.jp/cstp/english/society5_0/index.html (in English)

⁴ CPS stands for Cyber-Physical System.

the AI-centered systems will be interoperable to pursue human happiness beyond the barriers of industry and national borders.

Table 1: Key characteristics and issues of Society 5.0

	Society 4.0 and earlier	Society 5.0	
	Characteristics	Characteristics	Issues
Daily life and digital technology	Physical space and cyber space are separated.	Cyber space is integrated to physical space.	It is very complex. Privacy, security, fairness, democracy, and environmental threats may arise.
Object of trust	Tangible (people, things)	Intangible	Who and how to evaluate the reliability of data and algorithms
Acquirable data	Limited	Greater scale, scope and variety	Need to consider both the reliability of the data to be obtained and effective governance using the obtained data.
Decision-making actor	Only Humans	Expanding the impact of AI systems	Limited foreseeability
System conditions	Stable	Fluid	Fast-changing, where things change quickly.
Predictability & controllability	Relatively predictable and controllable.	More areas become unpredictable and uncontrollable.	Uncertainty always exists.
Responsible actor	Easily identified	Not easily identified	No one can take risks because the entity responsible when problems occur is unclear.
Concentration of control/power	Predisposed to concentration	More predisposed to concentration	The emergence of entities with certain dominant positions makes governance by other entities difficult.
Geographical relationships	Local OR global	Local AND global	Cooperation in international governance is essential.
The "Goals" of the system		There are multiple "goals" that need to be balanced.	Need to constantly carry out ongoing analysis of the social conditions and define the "goals."

Design and Implementation of Agile Governance

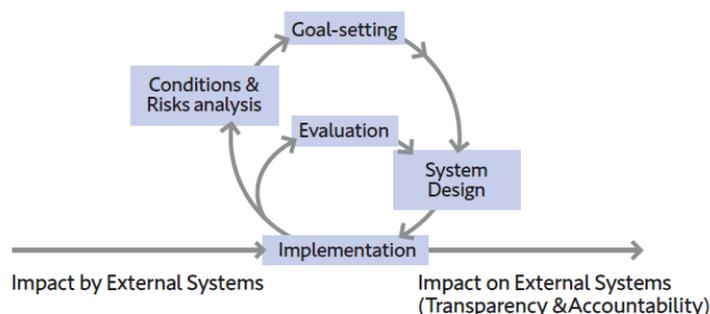
In Society 5.0, CPS which form our social infrastructures will undergo changes in

complex and rapid manner, making them difficult to predict and control. That being said, the "goals" themselves also will constantly change along with changes in social conditions. In light of these changes, the governance model for Society 5.0 must be one where solutions are constantly revised to ensure their optimality based on environment, technology and goals that constantly change. For this reason, the idea of agile governance has been proposed under such conditions.

Agile governance is a multi-stakeholder governance model that continuously and rapidly run cycles of "conditions and risks analysis," "goal setting," "system design," "operations," "evaluation," and "improvements" (see Figure 1). In the first "conditions and risks analysis" phase, governing actor should constantly analyze the external conditions, changes to these conditions, and the risk landscapes that result from these conditions. Then, in the "goal setting" phase, the goals should be constantly reviewed in accordance with changes in the external conditions and technological impact. Based on the "conditions and risks analysis" and "goal setting," the "governance system design" involves the design of technological systems, organizational systems, and their applicable rules. This is followed by "operation of the governance system," including continuous monitoring of the status of system operation based on real-time data and appropriate disclosure to stakeholders.

In the "evaluation" phase, the governing actor evaluates whether the initially defined goals have been accomplished. The system is "re-designed" if these goals are not being met (inner, circular cycle). At the same time, continuous analysis should be performed on whether there have been any changes in the conditions and risk landscape, and if there have been, whether these changes necessitate revisions to its goals (outer, circular cycle).

Figure 1: Basic concept of agile governance⁵

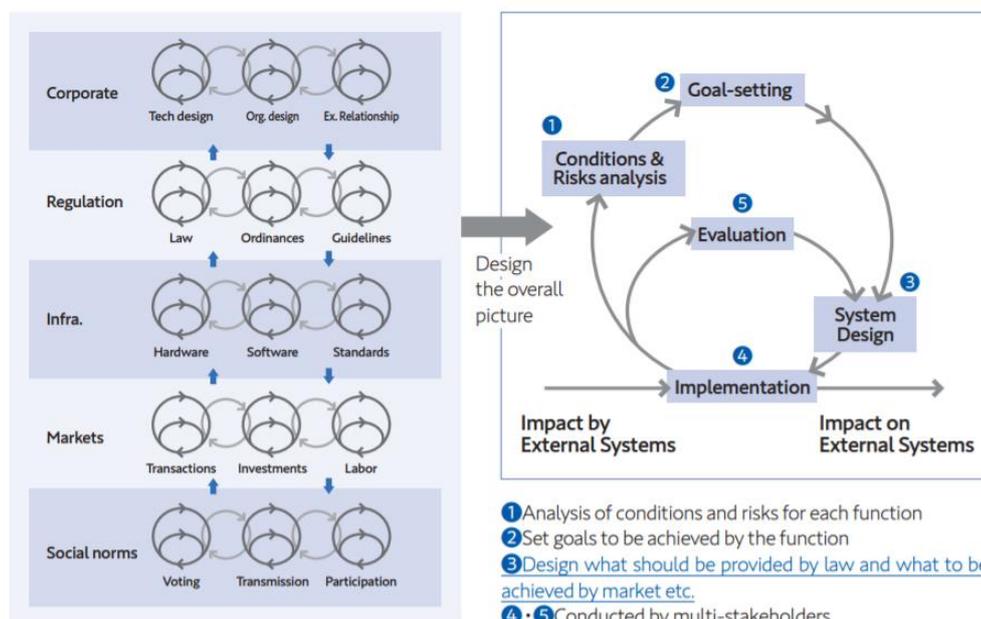


⁵ Excerpt from "[Governance Innovation ver.2: A Guide to Designing and Implementing Agile Governance](#)" (p.8, Japanese ver.)
 Excerpt from "[Governance Innovation ver.2: A Guide to Designing and Implementing Agile Governance](#)" (p.9, English ver.)

In order to implement agile governance of products and services and the systems, it is essential that businesses participate in governance. In addition to participation in governance, it is also important to encourage businesses to be accountable to their stakeholders (comply and explain). Designing appropriate incentive is an important factor in promoting corporate participation in governance and accountability. One effective idea to design incentive for businesses is for the relevant authorities to improve disclosure systems and provide compliance guidelines, thereby creating incentives for businesses to actively participate in agile governance. It is also important to design a comprehensive corporate sanction regimes in a way that focus on risk management and system improvement.

The key to realizing Society 5.0 is not only the implementation of agile governance by businesses. In the future, it will be necessary to implement agile governance in the governance systems of each major functions such as regulations, infrastructure, markets, and social norms, and at the same time, to clarify the overall picture by organizing the relationships among the governance systems for individual functions. Once the blueprint for governance is designed, the implementation of agile governance for a comprehensive governance system (“governance of governance”) will be indispensable for the realization of Society 5.0 (see Figure 2).

Figure 2: Concept of “Governance of Governance”⁶



⁶ Excerpt from “[Governance Innovation ver.2: A Guide to Designing and Implementing Agile Governance](#)” (p.16, Japanese ver.)
 Excerpt from “[Governance Innovation ver.2: A Guide to Designing and Implementing Agile Governance](#)” (p.15, English ver.)

3. AI Governance in Japan

In the second part, Mr. Sugano introduced the outline of "AI Governance in Japan Ver. 1.0 (Interim Report by Expert Group on Architecture for AI Principles to be Practiced)"⁷ released on January 15, 2021 by METI.⁸

Domestic and overseas trends in AI Governance

As agreed upon by several countries in the OECD "Council Recommendation on AI" and the "G20 AI Principles," a global consensus has been reached on various principles for the social implementation of AI. As a result, domestic and international discussions have shifted from AI principles to governance for the social implementation of AI principles.

At present, a risk-based approach⁹ for AI governance design is very much an international common ground. With countries such as the EU and the U.S., as well as the U.S.-Japan Business Council, taking a positive view of regulatory design based on the risk-based approach. On the other hand, there are differences in the approach to the number of risk levels and the classification of risks in terms of specific risk assessment methods and criteria, and it cannot be said that countries, regions, and stakeholders have necessarily reached a consensus.

The METI has organized the AI governance structure as a layered structure (see Figure 3). Within this layered structure, the technology-neutral goals to be ultimately protected by AI governance are defined by AI principles, high-level guidance from AI principles to implementation, and practical guidance. Horizontal intermediate rules, such as legally binding regulations, legally non-binding guidelines, and international standards, have been established to achieve the ultimate goal. In addition, rules focused on specific targets are set separately, and these regulations and guidelines are monitored and enforced.

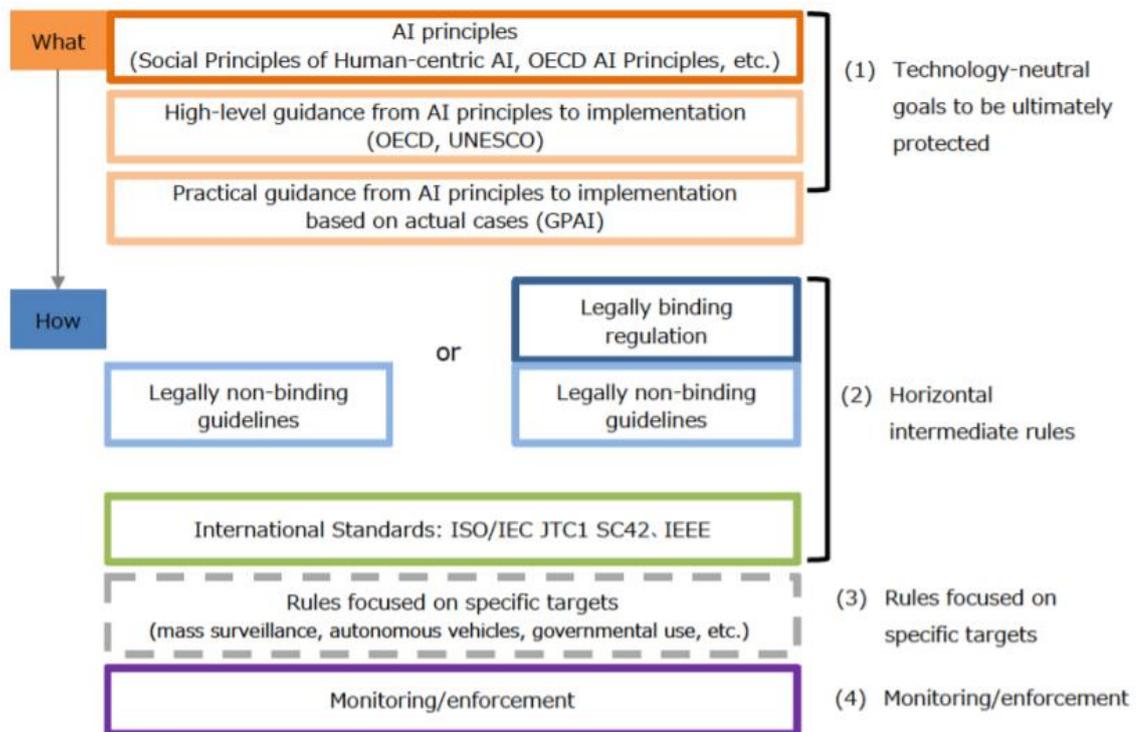
Based on this layered structure, international harmonization and alignment between layers are important when discussing AI governance.

⁷ <https://www.meti.go.jp/press/2020/01/20210115003/20210115003-1.pdf> (in Japanese)
<https://www.meti.go.jp/press/2020/01/20210115003/20210115003-3.pdf> (in English)

⁸ <https://www.meti.go.jp/press/2020/01/20210115003/20210115003.html> (in Japanese)
https://www.meti.go.jp/english/press/2021/0115_006.html (in English)

⁹ Idea that the degree of regulatory intervention should be proportionate to the impact of risks

Figure 3: Architecture of AI governance¹⁰



Ideal approaches and future issues to AI Governance in Japan

In Japan, laws and regulations face difficulties in keeping up with the speed and complexity of social changes, including AI innovation. To address such conflicting problems, it is considered to change the nature of regulations from rule-based ones to goal-based ones. In this policy shift, there is a possibility of a gap between goal-based regulations and operations. However, the gap is expected to be overcome by developing non-binding intermediate guidelines and standards.

In Japan, there are various opinions on the regulation of AI applications. The industry as a whole is generally in agreement that some form of regulation, including soft law, is necessary for AI applications. On the other hand, there are those who agree in theory with introduction of regulations for high-risk AI applications, but argue that the scope of regulation should be carefully defined.

Statistical research of consumers conducted by Fujitsu Research Institute revealed that consumers feel insecure about AI. Meanwhile, consumers are positive and place relatively high expectations about AI on areas where use of services can be easily imagined, such as "healthcare/ nursing care" and "automatic translation."

¹⁰ Excerpt from "[AI Governance in Japan Ver.1.0 \(Interim Report\)](#)" (p.9, Japanese ver.)
 Excerpt from "[AI Governance in Japan Ver.1.0 \(Interim Report\)](#)" (p.9, English ver.)

Based on the above, the interim report suggests that there is no need for legally binding horizontal regulations in Japan at the moment, but it also suggests the need for legally non-binding corporate governance guidelines to support goal-based governance.

The report also suggests that the discussion on AI governance is still in its infancy and that it would be desirable to discuss and examine the following issues in the future.

- ✓ Ensuring incentives to use the non-binding intermediate guideline
- ✓ Introduction of guidance on the use of AI by government
- ✓ Harmonization with other countries' governance and coordination between policies and standards
- ✓ Monitoring and enforcement, including monitoring the status of guideline usage

4. Discussion points in the question and answer session

In the 9th session, discussions were held on the AI governance and the contents of studies on the area of governance innovation in Japan. The following questions and answers were raised based on the topics discussed.

Notes on the relationship between the entities of each action of agile governance

- ✓ System developers and operators need to maintain their independence from AI service providers.
- ✓ The evaluation should be conducted by multi-stakeholders, including internal audits and external audits of outsourced companies. Transparency is also important for evaluation.
 - If a person in the center of a company covers up internal fraud, transparency will not be ensured, so it is necessary to consider the operation of a sanction system to some extent. In addition, it is necessary to design a sanction system in which the risk of disclosing information inside and outside the company is smaller than the risk of covering up internal fraud, such as by enhancing the whistleblowing system.

Designing incentives for corporate participation in agile governance

- ✓ It is necessary to consider the design of positive incentives, such as punitive incentives where there is a negative impact if governance is not implemented, or the creation of mechanisms where the implementation of governance is evaluated by investors.

Relationship to other government documents

- ✓ The “Principles of Human-centric AI society”¹¹, formulated under the Council for Science, Technology and Innovation, is considered to be one goal.
- ✓ METI is expected to prepare the contents of the report by focusing on the management of organizations and companies and the creation of rules, and how the report can be used by businesses as a whole.
- ✓ The METI report can be used to visualize the layers of each report and complement each other.

The relevance of clients and AI vendors on risk assessment

- ✓ The current situation is that AI vendors have not been able to communicate the necessity of risk assessment to clients who actually provide AI services.
- ✓ The client side also does not require risk assessment at the moment.
- ✓ It is important to have good governance for both the AI vendors and the client side of the services.
- ✓ Reports related to AI released so far have been for so-called AI experts, but in future, it is necessary to consider how to make the general public, who are not experts in the field, aware of the risks associated with AI in order to increase their literacy.
- ✓ It is difficult to assign responsibility for risks that arise from AI service users. We believe that AI service providers should be held responsible when they do not take measures to avoid risks that can be assumed, or when risks can be assumed by were not assumed.

Analytical entity for risk assessment

- ✓ With regard to the subject of risk assessment, although it is desirable for each company to proactively conduct risk assessment after a goal has been set, it is difficult to leave all the work to businesses, so the government agencies are considering providing check lists, etc.
- ✓ Checklists, etc., should remain abstract to some extent, but the government agencies should provide detailed examples, so that the companies should refer to them in their risk assessment.

Precedents for risk assessment of AI services in Japan

- ✓ Specific methods of risk assessment have not yet been studied at the moment, and case studies need to be collected.

¹¹ <https://www.cas.go.jp/jp/seisaku/jinkouchinou/pdf/aigensoku.pdf> (in Japanese)
<https://www.cas.go.jp/jp/seisaku/jinkouchinou/pdf/humancentricai.pdf> (in English)

- ✓ Unlike overseas, the Japanese government is aware that checklists are more required as a risk assessment method. It is necessary for the government to strike a balance and provide hints such as industry-specific risk assessment methods.
- ✓ There is also the idea of clearly stating what issues and items were considered important when the risk assessment was outsourced.

Handling diversity and inclusiveness in governance innovation

- ✓ It is necessary to set goals that the system needs to be designed to allow for diversity and inclusiveness.
- ✓ Diversity and inclusiveness will be a point that needs to be taken into account in corporate governance as well in the future.

We will continue to discuss AI governance in Japan and abroad through this study group.

Written by Ayuto Makiguchi
Translated by Michiko Shimizu

<Outline of the 9th Session of the Study Group>

Date & Time: Thursday, February 18, 2021, 17:00-19:00 (Zoom)

Agenda:

- Topical presentations:
 - "Governance Innovation ver.2: A Guide to Designing and Implementing Agile Governance" provided by Mr. Hiroki Habuka (Information Economy Division, Commerce and Information Policy Bureau, Ministry of Economy, Trade and Industry)
 - "AI Governance in Japan Ver. 1.0 (Interim Report by Expert Group on Architecture for AI Principles to be Practiced)" provided by Mr. Hirofumi Sugano (Information Economy Division, Commerce and Information Policy Bureau, Ministry of Economy, Trade and Industry)
- Question and answer session / discussion