

**Study Group ‘AI governance and its Evaluation’
Report on the Session #13**

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 13th session (April 27, 2021), Mr. Ichirou Akimoto of the World Economic Forum (WEF¹) & NEC Corporation (NEC), and Mr. Toshikazu Imada of Sony Group Corporation (Sony) presented topics on the theme of practices on AI governance in businesses.

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

2. Case study on AI governance (WEF and NEC)

First, Mr. Akimoto presented a topic titled "Efforts to Develop a Framework for Responsible Limits on Facial Recognition Technology".

➤ **WEF's AI-related initiatives**

C4IR², established in San Francisco in March 2017, is an institution with centers around the world that promotes the creation and demonstration of global rules to harness technology and solve social problems. In particular, C4IR is implementing regulatory reforms for the areas of "Data for Common Purpose Initiative," "IoT, Robotics, and Smart Cities," "Artificial Intelligence and Machine Learning," and "Blockchain and Distributed Ledger Technologies". In terms of AI, not much has been done in Japan, but teams in San Francisco and France have been working on AI-related activities.

Among C4IR's initiatives, there is a wide range of initiatives related to AI that focus on areas where rule formation has not been promoted internationally. These efforts are focused on three strategic areas: "Enabling Frameworks," "High-Risk Use Cases," and "Leapfrog Opportunities." The topic of this session, "Responsible Limits on Facial

¹ WEF: World Economic Forum

² C4IR: Centre for the Fourth Industrial Revolution

Recognition Technology" (hereafter referred to as the "Facial Recognition Technology Guidelines"), is in the context of "High-Risk Use Cases". This guideline focuses on walk-through management in airports, among other areas of facial recognition technology. This is because privacy is an issue that has been raised as airport security issues have become more important since 9/11³, and this guideline was drafted to address the privacy issues that arise in the context of security measures.

The Facial Recognition Technology Guidelines have been created in a step-by-step scheme, with the first and second versions of the white paper being published. The partners in the creation of the first version of the guidelines were Charles de Gaulle Airport in France, SNCF⁴, AFNOR⁵, and other companies. However, due to the global outbreak of COVID-19, the use case demonstration at an airport in France that was to have been conducted was abandoned. For this reason, Narita Airport, which had already decided to introduce facial recognition walk-through management in accordance with the guidelines of the Ministry of Land, Infrastructure, Transport and Tourism (MLIT), and NEC, which is providing the facial recognition technology, became partners in the verification experiment.

➤ **AI governance development by WEF**

The WEF has adopted a flow management style process (see Table 1) for the development of Facial Recognition Technology Guidelines in order to develop multi-stakeholder, consensus-building guidelines.

Table 1: Process for Developing Guidelines for Facial Recognition Technology

	Phase	Activities
1	Scoping phase	<ul style="list-style-type: none"> • Design the project plan. • Identify and engage with potential pilot partners.
2	Drafting phase	<ul style="list-style-type: none"> • Draft a white paper.
3	Testing phase	<ul style="list-style-type: none"> • Test the implementability of the white paper created in the Drafting phase on real cases. • Revisd the white paper based on the validation results.
4	Deployment phase	<ul style="list-style-type: none"> • Deploy the guidelines developed by using webinars, ISO, and IEEE.

³ The September 11, 2001 terrorist attacks on the United States

⁴ SNCF (Société Nationale des Chemins de fer Français) : National Company of French Railways

⁵ AFNOR (Association Française de Normalisation) : French Standardization Association

In the current WEF, those who can contribute to the rule-making process take the lead in implementing the four phases mentioned above and formulating guidelines. If Japan wants to be involved in global rule making in the future, it is necessary for Japan to actively raise its hand and contribute to the process.

The methodology for the development of the framework consists of the following steps: 1. Definition of the events to be covered (Define step), 2. Best practices (Design step), 3. Check items called Assessment Questionnaires (Assess step), 4. Audit conducted by a third party organization (Validate step). Among these, Audit is a check by a third party organization to ensure that AI and system implementations are incorporated correctly, and its importance of this check differs greatly between Japan and other countries. In Japan, the Audit process is not so important, and if the system operator acknowledges that the system is working properly, the operation is approved, but in other countries, the Audit must be conducted by a third party. No audit has been conducted on the The Facial Recognition Technology Guidelines developed this time, only the framework of what a third party organization will do upon Audit was announced.

➤ **Testing phase example (Narita Airport use case)**

NEC and Narita Airport, which had already decided to demonstrate a facial recognition system for the Olympics,⁶ participated in the Testing phase of the guideline drafting flow to verify the white paper. As a result of this verification, the check items described in the white paper had to be revised. This is because the check items were not designed for a walk-through management case like Narita Airport at the time of the Drafting phase, and there were several items that became unnecessary when the actual verification was conducted at Narita Airport.⁷ Therefore, the white paper at the beginning of the Testing phase for Narita Airport had 63 questionnaires in 11 categories, but at the end of the Testing phase, it was revised and published with 61 questionnaires in 10 categories.

The guidelines are available in three languages (Japanese, English, and French), and will proceed to the Deployment phase using a webinar by the WEF.

⁶ <https://www.narita-airport.jp/jp/faceexpress/> (in Japanese)
<https://www.narita-airport.jp/en/faceexpress/> (in English)
<https://jpn.nec.com/safercities/transportation/fasttravel/narita.html>

⁷ The case study of Narita Airport deals with walk-through management, but there are still few examples of this in overseas airports. In many overseas countries, the model is that the acquired photos are retained by the airport, and the list of questionnaires was initially based on this model, and the right to delete photos was also included in the guidelines. However, in the Narita Airport model, the airport can only keep the photos for one day, and they cannot be taken out of the airport, so the user's right to delete the photos does not need to be considered. As a result, there were cases where questionnaires listed in the white paper were not valid in the Narita Airport model, leading to a reduction in the number of categories and items.

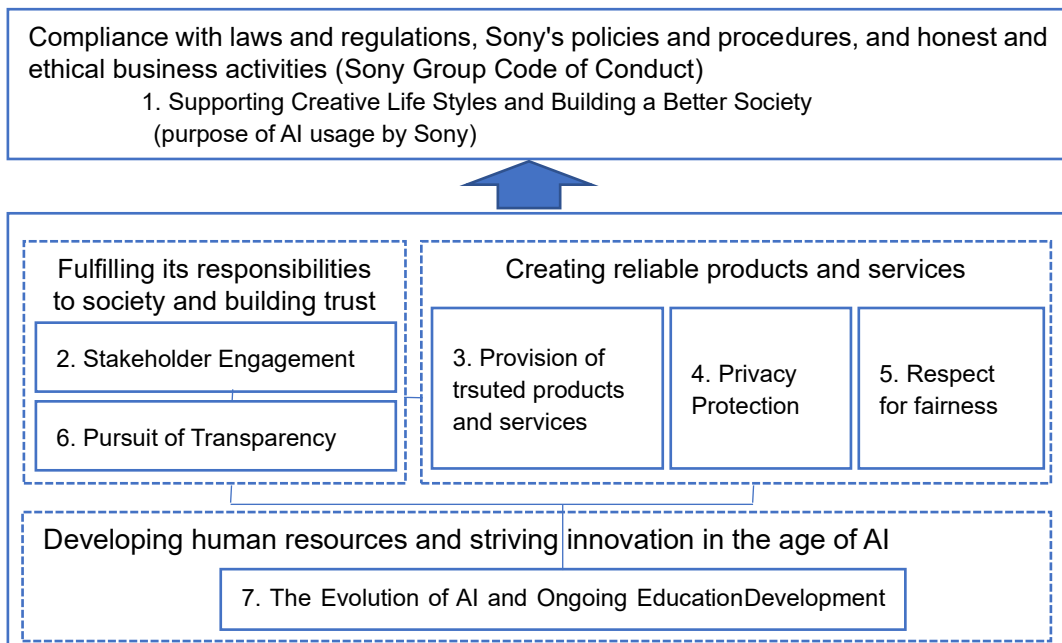
3. Case study on AI governance (Sony)

Next, Mr. Imada gave a presentation titled "Sony Group AI Ethics Activity".

➤ Operation of Sony Group AI Ethics Guidelines

The "Sony Group AI Ethics Guidelines" (hereafter "AI Guidelines") were released in September 2018 and updated in March 2019 in response to Sony's Purpose & Values. The AI Guidelines are based on Sony's ideal state based on its Founding Prospectus, Mission, Vision, and Sony Group Code of Conduct, and are consistent with external AI-related guidelines and other reference materials. The AI Guidelines are structured based on Sony Code of Conduct, which states, "Compliance with laws and regulations, Sony's rules and policies, and honest and ethical business activities, and defines the purpose of AI usage by Sony as " Supporting Creative Life Styles and Building a Better Society." In working to fulfill its purpose, the company is committed to "fulfilling its responsibilities to society and building trust" as well as "providing trusted services" and "developing human resources and striving innovation in the age of AI" (see Figure 1).

Figure 1: Principles of the Guideline

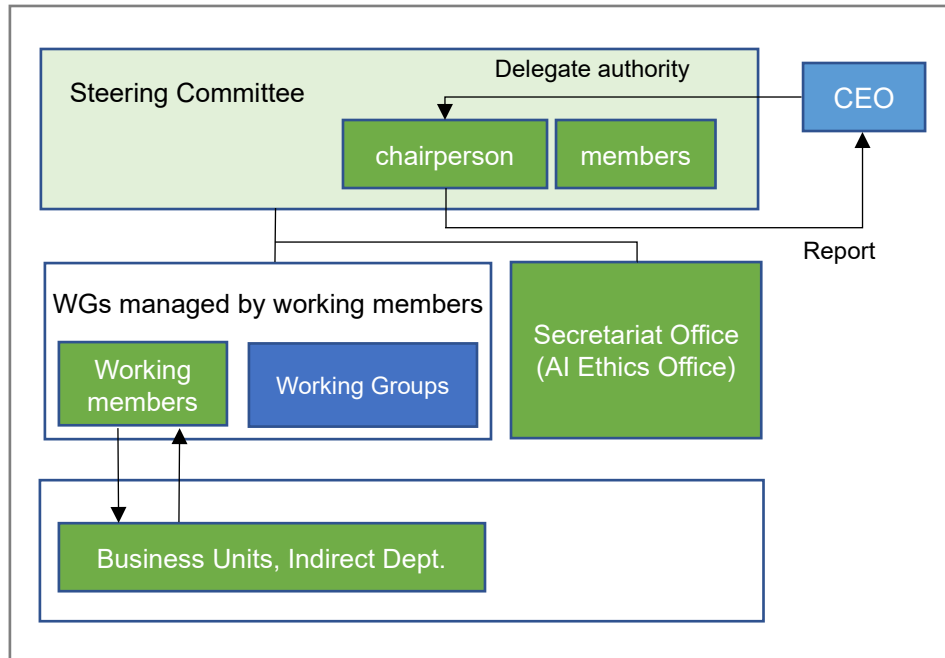


AI governance at Sony is led by the AI Ethics Committee. The AI Ethics Committee is composed of three layers (see Figure 2).

The Steering Committee consists of a chairperson appointed by the CEO and members appointed by the chairperson, with the CEO delegating authority to the chairperson, who reports to the CEO. Under the Steering Committee, there are working members and

working groups⁸, and measures are implemented for each business unit and indirect departments.

Figure 2: Composition of the AI Ethics Committee



One of the main AI governance initiatives within Sony is education and enlightenment. First of all, Sony is distributing an eLearning course for Group employees on what AI ethics are, an outline of Sony Group AI Ethics Guidelines, and an outline of data compliance, in an effort to establish AI ethics. In addition, the company shares and disseminates information through its internal portal and holds lectures by experts in various fields to raise awareness, and also disseminates information through its corporate website⁹ and external collaborations.

➤ **Sony's Assessment**

Sony's assessments are based on the concept of AI Ethics by Design, identifies the risk of social criticism, lawsuits, and brand damage due to incidents that violate the principles set forth in the guidelines at an early stage, and provides a risk response plan and implementation. In addition, Sony's AI ethics assessment process has been developed along with internal documents and assessment tools that stipulate compliance with the AI Ethics Guidelines in the commercialization process of Sony's electronics products

⁸ Solve actual problems.

⁹ https://www.sony.com/ja/SonyInfo/sony_ai/guidelines.html (in Japanese)
https://www.sony.com/en/SonyInfo/sony_ai/guidelines.html (in English)

and services. These ensure that AI ethics assessments are made in accordance with the policies of the Sony Group AI Ethics Guidelines in conjunction with Sony's existing rules and assessments.

The above-mentioned initiatives started with Sony's electronics business, and will be expanded to include financial and entertainment businesses in the future.

4. Discussion points in Question & Answer session

In the 13th session, practices on AI governance in businesses were discussed. Based on the topics presented, the following questions and answers were raised.

Actual operation of the framework created by the WEF

- ✓ Many of the check items in the white paper are not for vendors, but for actual users. Therefore, the checklist is based on the points to be considered when an airport actually introduces a facial recognition system.
- ✓ Generally, overseas, an audit by a third-party organization is conducted for the formulation of a framework. However, in this case, due to the impact of COVID-19, no audit by a third-party organization has been conducted, and the report only describes how the audit will be conducted. Therefore, this initiative is only at the stage where an Audit mechanism for the framework has been created. There are many significant differences between the model of facial recognition implemented at Narita Airport and the models currently in use at overseas airports, and it is expected that the Audit mechanism will be brushed up through the release of the second and third versions in the future.

Scope of the assessment

- ✓ As for the perspectives to be dealt with in the WEF check items, the scope will be determined in the Scoping phase and actually materialized in the Drafting phase. As for the technical aspects, the indicators set by NIST¹⁰, which conducted the verification of NEC's AI, are relevant.
- ✓ Since Sony is currently unable to conduct assessments for all AI ethics assessment subjects, it is in the process of prioritizing them based on appropriate risk assessment and implementing them sequentially on a pilot basis.

Governance from the perspective of an AI developer at Sony

- ✓ It is necessary to present to users what the risks are in the form of a Transparency Report. It is also necessary to update the report when updates are made to AI as

¹⁰ NIST: National Institute of Standards and Technology

well.

- ✓ As re-training, data addition, and updating of AI products are made, the idea is to take them through the planning phase of the commercialization process, where they are monitored and assessed.
- ✓ Audit of developed products/services is still an issue.

Governance from the perspective of an AI user at Sony

- ✓ In the case of using other companies' recognizers, etc., if there is something that can be tested as a final product, the policy is to test it and then if necessary, provide feedback to the vendor providing the product and request improvements.
- ✓ If the above is difficult, the company intends to explain the risks of the product to customers, including the limitations of AI functions.
- ✓ Discussions have begun within Sony on how to counter the potential risks posed by the AI currently in use, which could result in deviations from internal governance policies.

Key Considerations in Developing Sony AI Guidelines

- ✓ While focusing on AI ethics, the company valued the idea of not inhibiting the creation of innovation.
- ✓ Efforts are being made to ensure that the workload in the form of assessments does not become too heavy.
- ✓ It has been communicated within the company that the establishment of the guidelines does not mean that existing activities will be restricted, but rather that it is a new business opportunity.
- ✓ It is also important to note that employees should be aware of the growing social demand for AI ethics.

AI Regulation¹¹ prepared by the EU

- ✓ Since the European trend is likely to propagate worldwide, Japan should also make preparations for examining the regulatory content in stages and make efforts to cover them, even if it does not incorporate all of them.
- ✓ The key is how to reduce proposed AI regulations for the sake of industry and innovation while using one aspect of the regulations without over-regulation. For this purpose, the focus will be on whether we can propose persuasive mitigation measures by accumulating case studies.

¹¹ https://ec.europa.eu/commission/presscorner/detail/en/ip_21_1682

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 13th Session of the Study Group>

Date & Time: Tuesday, April 27, 2021, 17:00-19:00 (Zoom)

Agenda:

- Topical presentations:
 - " Efforts to Develop a Framework for Responsible Limits on Facial Recognition Technology" provided by Mr. Ichirou Akimoto (the World Economic Forum / NEC Corporation)
 - "Sony Group AI Ethics Activity" provided by Mr. Toshikazu Imada (Sony Group Corporation)
- Question and answer session / discussion