Study Group 'AI governance and its Evaluation'
Report on the Session #1(Phase Ⅱ)

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 to help build trustworthy AI systems, and the phase Ⅱ began in September 2021.

In the 1st session (Sep. 30, 2021), the study group chair Arisa Ema, Associate Professor at the Institute for Future Initiatives of the University of Tokyo, and also the board member of JDLA, presented the topic on "Who Manages and Evaluates AI?" This report is a reconstruction of the presentation and the discussions of the study group participants.

2. Who Manages and evaluates AI?

Associate Professor Arisa Ema, spoke on "Who Manages and Evaluates AI?" The presentation aims to introduce the background and contents of the report 'AI Governance Ecosystem, Trusted AI with Industrial Structure'[1] while also laying out the policies for future activities as part of Phase II.

Who is Managing and Evaluating AI?

A characteristic of Japan's industrial structure is that it often forms long, multiparty supply chains or Business-to-Business-to-Consumer (B2B2C). With various players involved in AI management and evaluation, it is challenging for a single organisation to provide overall management and evaluation due to the AI changing as it learns. Therefore, it is necessary to build a system that manages and evaluates AI that is not run by a single company but forms an "AI governance ecosystem" that includes external environmental factors such as insurance, auditing, standardisation, guidelines, and third-party committees.

---

[1] https://www.jdla.org/en/en-about/en-studygroup/en-sg01/

Trends in the Discussion of AI Management and Evaluation

In recent discussions about AI, the phrase "principles to practice" is often used. It is also used as a subheading in the Institute of Electrical and Electronics Engineers (IEEE) 2019 publication "Ethically Aligned Design."

Various AI-related principles such as the 'Social Principles of Human-Centric AI' have been formulated in Japan. Many other organisations and companies are developing principles centred around AI governance as well. However, who is explicitly responsible for these principles, not to mention how and by whom they should be evaluated is unclear. Indeed, the exact path of this 'principle to practice' methodology, is still undergoing a period of flux worldwide.

A governance system needs to be established to discuss AI management and evaluation. In the report 'Social Principles of Human-Centric AI' it was stated that "It is always necessary to continue to update the content and defined purposes surrounding Human Potential, Social Systems, Industrial Structures, and Innovation Systems in line with social changes and technological developments. For that reason, it is necessary to have a system that can be implemented and in place for various stakeholders, including government, industry, universities, research institutions, and the general public. Then they will be able to work together on such matters as identifying issues, evaluating impacts, and making decisions on regulatory governance including rules, systems, standardization and codes of conduct[2]."

AI principles need not merely be created; these principles need to be put into practice through governance, and a feedback loop created that collates outcomes and moves principals into practices needs to be established. However, it is possible that where responsibility lies when considering governance may become unclear during this process. One of the aims of the AI Governance Ecosystem proposed in the report is to visualise the roles of each actor in the different phases of AI systems when it comes to issues such as development, utilisation and incident response.

The Role of the JDLA Study Group

Many people have probably seen a variety of principles and guidelines relating to AI ethics and governance created by various organisations both in Japan and overseas. However, it should be noted that large companies are leading the creation of these

---

[2] https://www.cas.go.jp/jp/seisaku/jinkouchinou/pdf/humancentricai.pdf

principles and procedures. But at the same time, there are high expectations for local governments and startups to be the driving force of AI development and AI utilisation both domestically and internationally.

It is also difficult for a single company or department to make AI governance a reality. As such, the area of AI governance must become a collaborative space rather than a competitive one. This is why industrial organizations such as the JDLA will be needed to promote innovation and provide assurance for using AI. These organizations will also bridge the gap between domestic and international stakeholders as well as between startup companies and industry partners.

Characteristics of Japan's Industrial Structure

Companies called 'platformers' acting as large Business-to-Consumer (B2C) companies exist. On the other hand, many companies in Japan are Business-to-Business (B2B) which means that supply chains and the distance between developers and service users can be lengthy. Because it is debatable whether the cause of an accident can be traced through lengthy supply chains it makes difficult to provide users with correct and intelligible information. It can also make clarifying who is responsible complex.

In recent years, flows such as Consumer-to-Consumer (C2C) and Consumer-to-Business (C2B) has created additional supply-chain complications. It is, of course, essential for service providers to prepare for AI governance, and improve user literacy. Over and above this, it is vital to have a system that guards against and prevents accidents. As no single company can adequately deal with all of the risks involved, an AI governance ecosystem that includes its external environment needs to be constructed.

AI Governance Ecosystem

In AI governance ecosystem, what the industry should do is described in the 'AI Service Provider/ Developer'. There are governance bodies for service providers in the external environment, that include services such as 'quality control', 'auditing', and 'accident investigation', among others. These are expected to cooperate with AI service providers and their regular duties. On the other hand, there are governance bodies for users of AI products in the form of 'consumer protection'.

Monitoring the movement of governance bodies that influence policy and the formation of rules in this area is also essential. Every industry needs to take the lead in forming acceptable practices that promote fairness and safety while not overregulating.

There also exists an expectation that there will be cooperation between industry and

environmental management body. As drones and robots have AI systems installed in them, it is essential to consider how data reliability should be ensured in both real space and cyberspace. Who will monitor it, security management and whether data fraud can be internally reported are also essential factors that must be considered. AI uses data to learn and then turns that data into algorithms to provide services but exactly what kind of data was used in the learning process can sometimes only be known internally. Therefore, creating a system for whistleblowers against technical fraud is an important point.
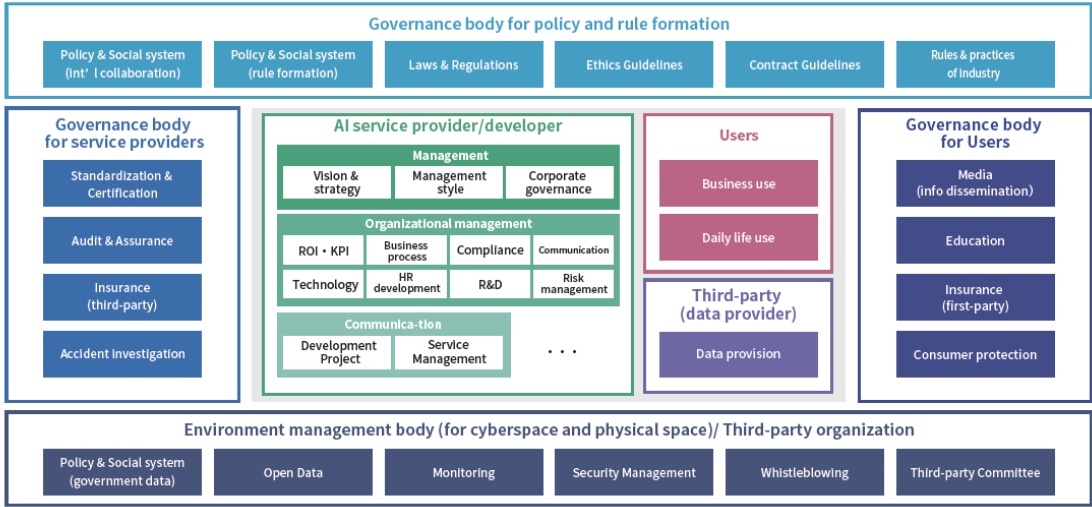

Fig. 1 AI Governance Ecosystem[3]

Because the parties involved in creating AI services will change throughout the various stages of development, utilisation, and problem-solving, separating them by stage and developing guidelines and AI governance ecosystems for each would be very useful.

For example, when developing AI, it is necessary to browse information from external organisations and link data from various sources. While AI is in use, having user supporting governance organisations that monitor and assist the user are essential. That way, in the event of a problem, arising, there is an out of organisation whistleblower who can cooperate with third parties and insurance.

Future Challenges for Japan

While the necessity of building AI governance ecosystem has been explained, it also needs to be mentioned that incentivisation is necessary to counteract financial costs and

---

[3] Excerpted from the Report "AI Governance Ecosystem Trusted AI with industrial structure "（Jul. 21, 2021）

promote active discussion between parties to form the ecosystem. An example of this incentivisation would be an organisation working with external stakeholders to build an AI governance certification framework for companies that provides the end-user with assurance. A single organisation cannot do this. Rather, it should be a collaborative effort between international and domestic organisations and certification bodies.

The idea of fairness concerning AI has also become a global problem. Due to Japan's long B2B2C supply chain, users may be unaware of the concerns upstream in development and vice versa. As such it is important to raise awareness of the topic of fairness by actively discussing these topics in the public domain. When discussing issues of fairness internationally the topics of statistical discrimination such as race and gender is widely discussed, however, in Japan human rights issues such as peer pressure and bullying are also prevalent and need to be considered. It should be noted that problematic issues vary from country to country and just because something is a problem in one country does not mean it will be so in Japan.

Furthermore, in Japan, there exist problems that cannot be dealt with using existing legal, financial auditing or insurance frameworks. Whether new laws are needed or whether problems should be dealt with in existing frameworks needs to be considered. For example, how physical or financial damages can be applied in a cyber environment is also an issue for AI that needs to be addressed.

Two Helpful Ideas

Here are two ideas for considering the ideal form of AI governance.

➤ The Collingridge Dilemma (1980)

It is difficult to predict the impact of technology before it is used in society, but it is difficult to control technology once it has become widespread.

This dilemma has always been difficult to evaluate from a technological perspective but it may be addressable with the agile development in AI. At the moment instead of performing a variety of social experiments, it needs to be remembered that we live in an experimental society and that the effects of any given technology need to be considered right from the development stage.

➤ Responsible Research and Innovation

Technological development usage and problems will have an impact and in recent years the concept of Ethical, Legal and Social Implications (ELSI) has been discussed for information technology both within companies and within policy circles and needs to

continue.   ELSI originated in the field of life science in 1990 around human genome projects. However, it is not currently used in Europe or the United States. Instead, Responsible Research and Innovation or RRI has become an important part of their research and development methodology since the early 2000s. Figure 2 shows the important aspects of RRI guidelines. These concepts have great relevance to current AI development, governance and guideline creation.
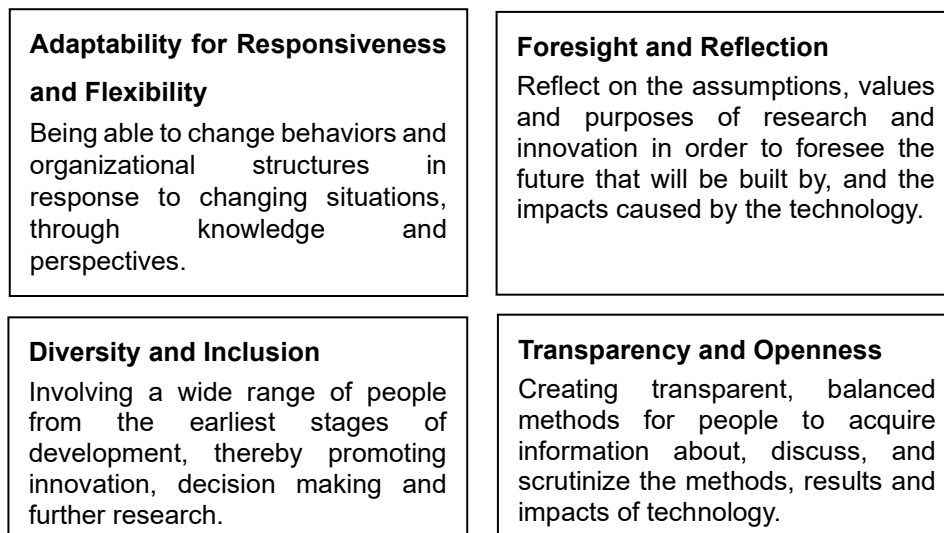
Adaptability for Responsiveness and Flexibility Being able to change behaviors and organizational structures in response to changing situations, through knowledge and perspectives.	**Foresight and Reflection** Reflect on the assumptions, values and purposes of research and innovation in order to foresee the future that will be built by, and the impacts caused by the technology.
Diversity and Inclusion Involving a wide range of people from the earliest stages of development, thereby promoting innovation, decision making and further research.	**Transparency and Openness** Creating transparent, balanced methods for people to acquire information about, discuss, and scrutinize the methods, results and impacts of technology.

Fig. 2 Responsible Research and Innovation[4]

3. **Comments from the Study Group Participants**

The following is a summary of the meeting Q&A session.

➢ Japan's future response to regulations such as those from the EU

✓ How to deal with the proposed European Union AI Act is currently being discussed by multiple government departments, companies and industry groups.

✓ For Japan to make a new law that are easy to use and adapt it must examine the current practices of those involved or the laws will become impractical. As such, a list of current practices needs to be made.

✓ Japan should collaborate with not only the EU but also with America, and South East Asia to ensure international cooperation.

➢ AI Governance Ecosystem Incentive Design

✓ Japan currently has little awareness of data fairness. Without incentives, there will be no opportunity to tackle this issue of equity.

✓ In the AI governance ecosystem, it is hard to see who should be responsible for

---

[4]Excerpted from materials of the meeting.

this incentivization in countries with B2B2C structures such as Japan. While the ecosystem must work with various groups and companies it must not impose responsibility on third parties.

- ➢ Product Liability for Self-Driving Cars
  - ✓ It is important to stop the AI learning process at a certain point, thereby allowing the specifications delivery point to occur and leaving an understandable user log.
  - ✓ As the issue of responsibility is extremely complicated it is necessary to think of responsibility in terms of an ecosystem. Even if it should prove difficult for a manufacturer to take full responsibility, what responsibility is taken needs to be explainable to insurers, consumers and certifiers in an easy to understand way. This again necessitates the creation of an ecosystem.
  - ✓ Liability for car accidents involving self-driving cars is split into two categories, civil liability and criminal liability. From an insurance perspective, covering the former is possible, but covering criminal liability is more difficult. An example of a countermeasure for this would be to have an accident investigation committee, much like in the aviation industry. This will allow for accidents to be investigated to prevent reoccurrence without pursuing responsibility or at least seeing responsibility as a secondary priority.
  - ✓ Some organisations such as the Japanese Electronics and Information Technology Industries Association (JEITA) and the National Traffic Safety and Environment Laboratory are considering the question of product liability for autonomous vehicles.

- ➢ Global Trends in AI Governance and Standardization.
  - ✓ The U.S and Europe are taking the lead in governance and have a large number of people working on it.
  - ✓ Even in China AI Governance is being discussed and the government has formulated governance rules.
  - ✓ In the Middle East, the UAE has been working on creating a network and breaking into Asia, Europe and the United States.
  - ✓ Thanks to ex-pats living in and returning from Europe and the United States, the conversation around governance has begun in Africa.
  - ✓ In Japan, the attendees at international seminars on the topic tend to be the same people. It is important to develop a skilled base of people with extensive knowledge who can make the process of creating and integrating into an

international framework simpler and faster.

✓ At present, it can be seen that there are many stakeholders in the issue of AI ethics but only a few people are discussing this topic.

✓ Europe has been working towards a hard law system whereas Japan has been working towards a soft law system which may lead to differing roles for each country. While working together is important it is also important not to impose upon each other's cultural identities.

➢ Data Bias

✓ B2B companies need to the purpose of their system and its complications, such as data bias and their other concerns, with each other.

✓ Many companies are unaware of data bias. If open data is used for their AI, there will still exist an element of bias. While it is difficult to examine all of the data used, internal checking mechanisms and mutual confirmation measures are required for any project. However, companies must first be aware of these issues in order to deal with them.

➢ Disseminating Information about AI Governance Ecosystems

✓ For an AI governance ecosystem to be formed it is important that we create an English version of our documents that can be easily found and disseminated. We must also find people who agree with us and not merely spread information.

✓ It is important to work to disseminate information, develop concepts and sure up support to deliver outcomes domestically as well as in the international arena.

The discussion of AI Governance domestically and internationally will continue through this study group.

（Written by: Yuki Kiyomi）

(Translated by David Shield)

<The 1st Session of the Study Group>

Date/time：September 30th(Thursday) 16:00-17:30（On Zoom）

Contents：

- Introduction to the JDLA's study group
- Discussion topic: 'Who manages and Evaluates AI?'
  Presented by JDLA study group chairperson Arisa Ema, Associate Professor at the Institute for Future Initiatives of the University of Tokyo, Board Member of JDLA
- Question and answer session / discussion