February 2023

JDLA

Study Group 'AI governance and its Evaluation'

Report on the Session #5 (Phase III)

1. Introduction

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 III began in 2022.

In the fifth session of Phase III (December 19, 2022), Yasukazu Hirata (Robust Intelligence Inc.) presented a topic on overseas case studies for the realization of AI governance, particularly on the trends surrounding AI governance in the United States.

2. "AI Governance and Quality Management in U.S. Companies" by Yasukazu Hirata (Robust Intelligence Inc.)

Trends in AI Governance in the U.S.

With the rapid expansion of AI use, risks during implementation have become more pronounced over the years (e.g., discriminatory predictions and drift, incorrect behavior of models, adversarial input, etc.). These are new issues not seen in conventional software. In fact, in Europe and the U.S., the leading AI countries, major incidents are starting to occur that can damage corporate value. For example, a large drift caused by the coronavirus pandemic occurred in the housing price prediction AI, resulting in a huge loss due to wrong purchase decisions (Zillow, U.S.), or a 20-fold difference in the amount of credit given to a husband and wife due to their gender despite the fact that they shared the same assets and there were no differences in typical feature values of the credit screening AI such as addresses, and the authorities requested an investigation (Apple Card).

In the U.S., in response to these large-scale incidents, various government agencies are developing various laws and regulations for the industry in the context of governance and risk management. And states such as New York City, Washington, D.C., and California are moving forward with more specific laws and regulations in response to the movement.

On the corporate side, major U.S. companies are building organizations to fulfill their social responsibility in AI and hiring specialists in this field.

Examples of specific responses by companies (cases for addressing fairness and model risk management)

Two examples of how companies are responding to these legislative and regulatory developments are presented below. First, in response to the enactment of New York City's Artificial Intelligence Hiring Bias Law (which requires companies that use hiring AI to conduct annual bias audits, publish audit results in job postings, and announce the use of hiring AI models; scheduled to take effect in the spring of 2023), major human resource services firms headquartered in New York State and others are already taking steps to address this issue. They are proactively verifying and auditing the fairness of the hiring AI models they use, adding not only race which is subject to the bias audits, but also gender and residential area.

The second is model risk management, which financial institutions in particular are working to strengthen, but there are many practical challenges to achieving the goal. In the case of one Fintech company, the compliance department made rules in accordance with the model risk guidelines and the data science department responded, but the compliance department did not understand the mathematical basis they presented, and as a result, the first and second line of defense model was not working well. As a solution, Robust Intelligence's Model Card, a quality evaluation report that visually summarizes the main points, or the so-called "AI Model Report Card," was incorporated into the operation. This enabled them to establish an appropriate governance system by managing necessary items on the model card and communicating with each department via this information. The lack of a common language between the compliance department and the DS department is a major issue in promoting governance, but the operational method using the model card is one of the useful tools in model risk management, which requires the implementation of cross-departmental measures.

AI Governance in Japan

As above, in Europe and the U.S., AI governance is seen as an urgent issue and risk management systems are being developed, and awareness of this issue seems to be increasing year by year.

Many Japanese companies, on the other hand, have yet to fully implement AI. In addition, data scientists are burdened with a great deal of administrative work, such as explanations and reports, that could be called "babysitting" of AI. At the stage of considering AI implementation, a culture that views AI utilization in an overly conservative manner has put the brakes on full-scale AI implementation by failing to use

AI in core operations and using only in low-risk operations.

Considering the current situation in Japan from the viewpoint of governance, the lack of AI governance, which indicates "what and to what extent can be done with AI," may be one of the reasons for the lack of progress in considering AI adoption. When considering AI governance in the future, it is important to take an "offensive" as well as a "defensive" perspective in order to promote the future use of AI by Japanese companies.

3. Main comments from the participants

Following the topic presentations, Kojin Oshiba, Co-Founder of Robust Intelligence, joined the discussion. The main discussion topics are as follows.

- **Cross-company functions for compliance response**
  - ✓ Since it is impossible to respond to various laws, regulations, and guidelines all within the management process of each individual company, it may be necessary to have a cross-company role for those areas where common items can be indicated. (NIST is trying to create playbooks and databases for those areas where confidentiality is not an issue.)
  - ✓ Looking at trends in cybersecurity and information security (SOC2, ISO27001, NIPPA, etc.) in the U.S., startup companies (e.g., Vanta, Drata) are emerging to take independent responsibility for areas that are difficult to handle by individual engineers in the field. Similar to this trend, it is expected that specialized vendors will take charge in the AI field in the future.
  - ✓ These specialized companies have channels not only for technical checks but also for audits and legal aspects, and they serve as a bridge to companies that need these services. This business model is distinct from that of platform providers, and while those providers may be responsible for technical aspects (e.g., automatic deletion of data) in the future, it is likely to coexist within the ecosystem as a coordinating function for the final regulatory and guideline-related aspects.

- **Driving force and motivation within the organization required for AI governance**
  - ✓ In the U.S., businesses themselves are becoming increasingly dependent on AI services, and it is likely that investment in AI governance is increasing along with investment in technology. In Japan, however, there are some difficulties in promoting AI governance in terms of difficulty in explaining the return on investment.
  - ✓ It is important for management to recognize the importance of governance and promote it from the top down. By doing so, it is highly possible to solve the common problem of not being able to promote governance because the issues are focused

on short-term ROI at the divisional level. In the case of the response to the NYC regulation, the need for compliance with regulations was first voiced at the board level, and then the CEO instructed the Chief Data Officer, who then formed an AI Ethics Committee, thereby establishing it as an internal activity. The strong impetus is coming from the high levels of the company that are not directly involved in AI.

- ✓ In the U.S., budgets are often set up in risk management and security departments; IT security departments are more accustomed to discussing ROI for AI governance than DS departments.
- ✓ It is necessary to show that governance is not only an organizational or policy discussion, but that measures fall into the aspects of maintaining model performance and improving UX, which are the responsibility of the data science team in the field. In this way, governance can be promoted in a way that responds to appeals from all parties involved: the long-term perspective of upper management, governance ROI discussions, and KPIs from the data science team's perspective.

- **Integration of external professional elements and internal knowledge**
  - ✓ In the U.S., there has been a movement to hire AI governance promotion personnel, but there have been scattered cases where these personnel have had difficulty convincing all relevant parties within the company. Perhaps a two-party team of personnel who can act as a bridge to external parties and those who are familiar with the company's internal affairs would work well.
  - ✓ When considering the application of governance to AI models in a company, it is often the case in the U.S. to first target one high-risk model since it is impossible to target all models at once. As the next step, when applying governance to the entire company, internal personnel who are familiar with the organization and its operations will play an important role in determining the order in which the governance should be horizontally deployed.

- **Roles and support expected from the government, universities, industry associations, etc.**
  - ✓ In Japan, where individual companies tend to be more cautious about introducing AI than in Europe and the U.S., the government guidelines may play a significant role as "offensive governance" to encourage the promotion of AI. For example, when the domain and use of AI is unclear in terms of "fairness," if the guidelines mention that item and describe the minimum requirements to be followed, it will serve as a safety rail and provide reassurance that AI can be used within those limits, thereby encouraging companies to adopt AI.

- ✓ Universities and industry associations such as JDLA have a major role to play in the area of human resource development. Raising awareness by addressing the importance of AI governance in training programs and certifications can help set the direction of necessary career paths and talent flow throughout the industry.
- ✓ Industry associations can also serve as a coordinator and a consolidator for companies to lobby the government from the industry. As a preliminary step, they are also expected to collect and present corporate governance and incident case studies.

The 5th Session (Phase III) of the Study Group
Date/Time: December 19th (Monday) 11:00-12:00 (On Zoom)
Contents:
- "AI Governance and Quality Management in U.S. Companies" by Yasukazu Hirata (Robust Intelligence Inc.)
- Questions & Discussion