※当翻訳は仮訳であり、正確には原文を参照してください。

※Please refer to the original text for accuracy

※仮訳

※Provisional Translation

Interim Report

February 4, 2025

AI Strategic Council / AI Institutional Study Group

Contents

Overview

AI Strategic Council /AI Institutional Study Group
Overviews of Interim Report

Since July 2024, seven meetings in total of the AI institutional study group[1] have been held. This report was created after discussions including interviews with total of 15 researchers, business operators, etc., and public comments.

Background

- While AI has the potential to greatly contribute to Japan's development, **various risks are becoming apparent**.
- There are many concerns over AI, pointing out **that its development and use are not progressing** compared to other countries.

▶ It is necessary to **ensure appropriateness** such as the transparency of AI, and **advance development and use of AI**.

Basic Concept

Promotion of innovation while mitigating risks (Ⅱ.3.)

- Promoting innovation including support to research and development, human resources development and expansion of data or computing resources
- Appropriate combination with laws and guidelines
- Use of the common principles such as the OECD AI principles and Hiroshima Process International Guiding Principles for Organizations Developing Advanced AI Systems and individual existing laws

International cooperation (Ⅱ.4.)

- Leading discussion toward formulation of AI governance
- Ensuring international consistency and interoperability

Specific Directions for Systems and Policies

Overall items (Ⅲ.1.)

Aiming to be "the friendliest country in the world for developing and implementing AI"

Prompt legislation is necessary
A system aiming to be a global model

- **Strengthening the government's function of a strategic leadership board and formulating AI strategy**
  - Strengthening function of AI strategic leadership board to oversee the whole picture
  - Formulating AI strategy (AI basic plan) to contribute to promoting safe and secure research and development, and use of AI.
- **Improving AI safety**
  - Establishing guidelines by the government (in accordance with the Hiroshima AI Process), and cooperation by business operators.
  - Conducting surveys and collecting information by the government, and guidance, advice and information provision, etc. to business operators and citizens

Use of AI by the Government (Ⅲ.2.)

- Appropriate AI government procurement and use, etc.

Use in infrastructure services (Ⅲ.3.)

- Response by each business act, etc.

1) "AI Strategic Council" was established under "the Integrated Innovation Strategy Promotion Council", which is chaired by the Chief Cabinet Secretary and consists of all cabinet ministers. "AI Institutional Study Group" was established under the Council.

2) After taking the above measures, legal systems should continue to be reviewed in order to mitigate future risks.

I.Preface

Since autumn 2022, generative AI has experienced a dramatic leap in performance, enabling it to produce natural conversations, write code, and create sophisticated videos by processing vast amounts of information. This advanced AI has the potential to replace tasks traditionally performed by humans and even produce outcomes that surpass human achievements. It is anticipated to find applications across all fields of human activity, significantly enhancing efficiency and convenience in various industries and everyday life and to become a role to supplement new science discovery and humans' creativity. In the future, it is expected to contribute significantly to improving quality of life and driving national economic development.

On the other hand, various risks with AI are becoming apparent, including sophisticated criminal behavior such as fake websites and voice scams, and information manipulation by using and spreading AI to create disinformation and misinformation. There are also concerns over national security risks by using AI in the development of CBRN (chemical, biological, radiological, and nuclear weapons) and cyberattacks as AI has an aspect of dual-use technology.

Under such circumstances, in the EU, the AI Act, a comprehensive regulation on AI, came into effect in August 2024. The AI Act adopts an approach in response to four levels of risks. It has introduced regulations prohibiting AI that threatens human's safety and fundamental rights from putting into markets or using, and imposing an obligation to conduct impact assessments and conformity assessments before putting a high-risk AI, which is likely to seriously harm the health and safety of humanity and democracy and the rule of law, into the market. In addition, the Act has established regulations imposing obligations of the transparency including creation of technological documents and disclosure of training data on providers of the general-purpose AI, and additionally imposing an implementation of model evaluation or a report obligation regarding incidents on providers of a general-purpose AI with systemic risks, such as those having computational volumes of training that is more than 10^{25} FLOPs.

In the United States, since July 2023, major AI development companies have announced a voluntary commitment to implement information sharing regarding risk management of AI, research regarding social risks possibly caused by AI system and investment in cyber security. Additionally, in October 2023, to mitigate national security risks, based on the Defense Production Act, an Executive Order was issued, instructing US companies who develop dual-use foundation models with which computational volumes of training are over 10^{26} FLOPs, to continuously submit information on activities related to models' training, development or manufacture to the government, but the order was revoked on January 20, 2025. Three days later, on January 23, a new executive order related to AI was announced, indicating a review of policies and regulations, which are formulated under the previous order, and to conduct an action plan. Additionally, in September 2024, State of California enacted the state law SB 942 to improve the transparency of contents generated by AI and the state

law AB 2013 to disclose the data used for training AI.

The G7 in 2023, which was under Japan's presidency, launched "Hiroshima AI Process[1]" to consider global rules related to generative AI, and established "Hiroshima Process International Guiding Principles for All AI Actors" and "Hiroshima Process International Code of Conduct for Organizations Developing Advanced AI Systems" (hereinafter referred to as "Hiroshima Process International Guiding Principles, etc.") aimed at realizing safe, secure, and trustworthy AI. Subsequently, Japan, as an outreach beyond the G7, launched the Hiroshima AI Process Friends Group, working on expanding the number of countries and regions that support the spirit of the Hiroshima AI Process. In 2024, Italy, as presidency of the G7, took over the Hiroshima AI Process, and discussions on a framework to monitor implementation of the international code of conduct have been taking place. In addition to these, there have been active discussions on AI governance in frameworks among multilateral countries including the United Nations, the Council of Europe, and the OECD.

In Japan, since May 2023, "TENTATIVE SUMMARY OF AI ISSUES" (AI Strategic Council, May 26, 2023) and "General Understanding on AI legal system" (AI Strategic Team, May 2024) have been completed, summarizing discussion points on AI and indicating general understanding on AI legal system, and "The Integrated Innovation Strategy 2024" (decision by the Cabinet, on June 4, 2024) establishes strategy aimed at strengthening competition in the field of AI and ensuring safety and security. Furthermore, "AI Guidelines for Business Ver 1.01" (Ministry of Internal Affairs and Communications, and Ministry of Economy, Trade and Industry on November 22, 2024) indicates an approach to develop, provide, and use AI in business activities, and for relevant ministries to cooperate in forming AI policies.

According to the result of a survey on awareness for AI risks and safety (Figure 1) in Japan, as low as 13% of respondents think that "they can use AI safely under current rules and laws," and 77% of them think that "AI needs to be regulated." Furthermore, they feel risky about "quality instability," "process black-boxing" and so on, and they also call on the government to "strengthen legal measures against misuse of AI or crimes using AI." In terms of dealing with risks of AI, while in countries, especially European countries and the U.S., discussions and considerations of the legal system related to AI are progressing, Japan has responded mainly by pursuing "soft laws" such as guidelines, and has not yet considered a legal system specifically for AI.

Based on the above situation, the AI Institutional Study Group was established under the AI Strategic Council. It holds interviews with various stakeholders including business operators, experts

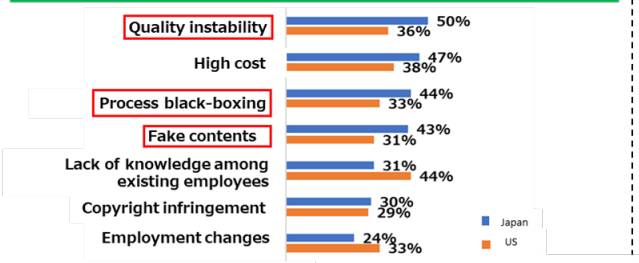
[1] Based on the result of the G7 Hiroshima Summit held in May 2023, this process was established in May 2023 to discuss generative AI, whose rapid development and expansion have become significant issues for the international community as a whole.

and local governments, and is considering how the AI legal system should be, including whether it is necessary or not. This consideration was based on the spirit of the Hiroshima AI process, and discussions were held based on four basic principles: “Balance between risk mitigation and the promotion of innovation,” “Design of flexible legal systems that can respond to the speed of technological and business change,” “International interoperability,” and “Proper procurement and use of AI by the government.” This report summarizes the result of considerations based on the interviews and discussions by the AI Institutional Study Group. Relevant ministries considered to interpret intellectual property laws, including copyright, in the process of development and use of AI. As for the use of AI from the perspective of national security, ministries and agencies related to the field are leading progressive discussions on considering AI specifically.

Survey on Awareness for AI Risks and Safety (International Comparison)

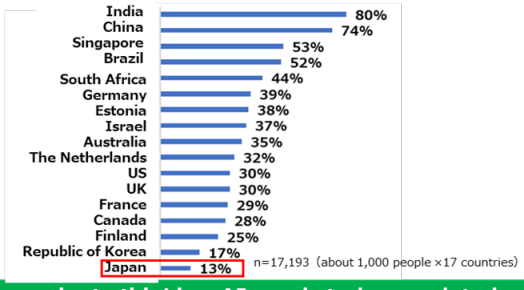
- Compared to the United States, in Japan, more companies are **concerned about "quality instability," "process black-boxing," and "fake contents,"** and **fewer companies take governance against AI risks**
- In Japan, only **13%** of respondents think **that "they can use AI safely under current rules and laws," the lowest** among the surveyed countries, and **77% of them think that "AI needs to be regulated,"** ahead of the United States, France and Germany.

Risks perceived regarding the use of generative AI



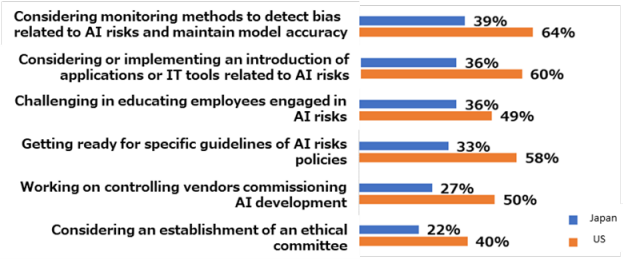
multiple responses, n=331 (Japan), n=1,014 (US)

Respondents thinking they can use AI safely under current rules and laws



n=17,193 (about 1,000 people ×17 countries)

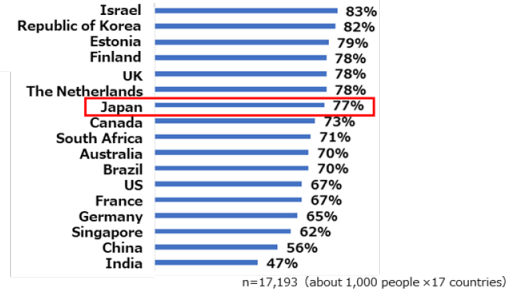
Status of governance measures for AI risks



multiple responses, only companies that are already working on or considering AI governance were asked, n=228 (Japan) , n=840 (US)

(Source) Prepared by the Cabinet Office based on the PwC Japan "2023 AI Predictions"

Respondents thinking AI needs to be regulated



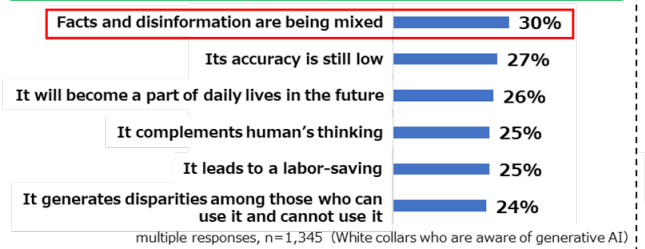
n=17,193 (about 1,000 people ×17 countries)

(Source) Prepared by the Cabinet Office based on the KPMG "Trust in Artificial Intelligence: A global study"

Survey on Awareness for Generative AI (Japan)

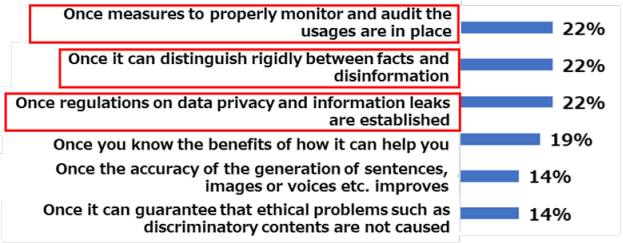
- As for the image of generative AI, 30% of respondents answered that **"facts and disinformation are being mixed,"** while in terms of what would make you inclined to use generative AI, 22% of respondents answered, **"measures of monitoring and auditing," "a rigid distinction between facts and disinformation" or "regulations for privacy etc."**
- The most common request for companies (66%) is **"to strengthen safety of customer data and privacy protection,"** while the most common request for government (66%) is **"to strengthen legal measures against misuse of AI or crimes using AI."**

The image of generative AI



multiple responses, n=1,345 (White collars who are aware of generative AI)

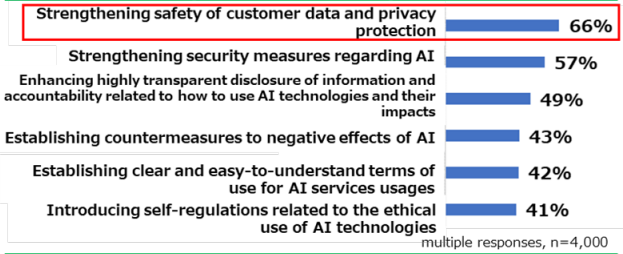
What would make you inclined to use generative AI



multiple responses,
n=620 (the willingness to use of generative AI services : respondents of "I cannot say either"-"I do not want to use them at all")

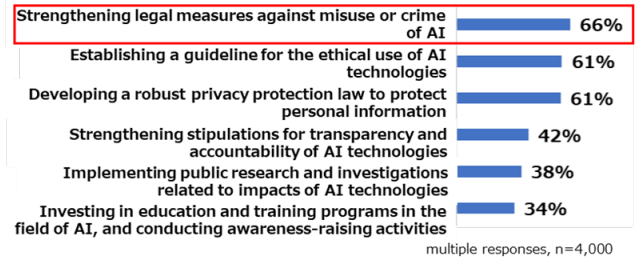
(Source) Prepared by the Cabinet Office based on the Cross Marketing "Survey on Generative AI (2023) the Actual State of Use and Awareness"

Requests for companies



multiple responses, n=4,000

Requests for government



multiple responses, n=4,000

(Source) Prepared by the Cabinet Office based on International University Global Communication Center "Innovation Nippon 2024 Generative AI and Japan"

II.Basic approach to legal systems

1.Recent developments in AI

There are various forms of AI taxonomy; for example, it can be classified into Specialized AI and General-purpose AI. Specialized AI is one that is specialized in processing specific tasks such as voice recognition, image recognition, or autonomous driving. General-purpose AI is one that is trained on larger amount of data than specialized AI, has high versatility and can process various tasks. Generative AI[2], whose potential has attracted a great deal of attention in recent years, generally belongs to general-purpose AI. It had been considered that general-purpose AI generally improve its performance as the number of parameters of training data and models increases, but, recently, general-purpose AI with high performance regardless of scales of training data has also appeared. Some people say that in the future, AGI (Artificial General Intelligence), which will have an ability to realize various tasks at the same level as humans, will appear. As described above, AI technology has been making remarkable progress in recent years. The definitions of terms related to AI, including “AI” and “business operators,” have been discussed internationally, and it is considered that these definitions will continue to change from now on, so it is also important to take these discussions into consideration.

2.Related actors

（1）Main actors

In this report, in the lifecycle from data collection or model development to AI system development, (including AI models), and finally use of AI services, we will describe on the premise of three main actors; AI developer, AI provider and AI user. (Figure 2)

First, AI developers shall be those who collect data, train models, and develop constructions of models' system infrastructure or input/output functions. Second, AI providers shall be those who incorporate AI into existing or new systems, and provide the AI systems so that it is ready for use in services, or those who implement everything from incorporating AI to providing AI service. Finally, AI users shall be those who incorporate AI systems implemented by others into their own services to use them as AI services, or those who use the provided AI services.

Furthermore, in addition to the three main stakeholders listed above, it is also necessary to be aware of the existence of various other stakeholders: those who provide training data, those who provide necessary resources for AI such as data centers, and those who conduct research.

---

[2] It is defined “a general term representing AI developed from an AI model that can generate texts, images, programs, etc.” in “AI Guidelines for Business Ver1.01”

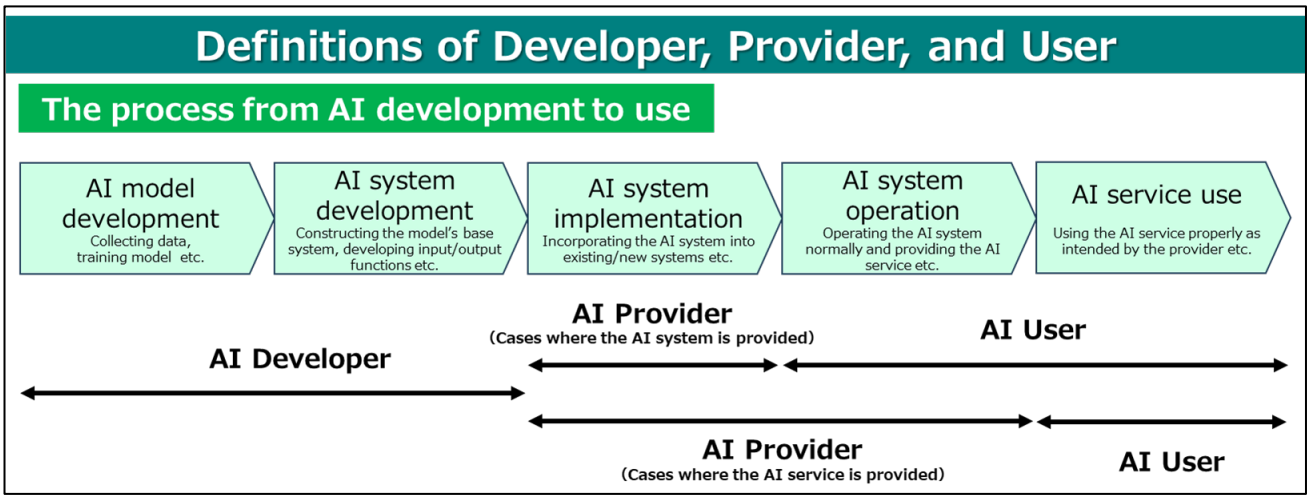


Figure 2 Positioning of each actor

（2）Overseas actors

Much of the generative AI used in Japan is provided by overseas business operators; it is not appropriate to exclude overseas business operators uniformly from the subject in considering AI-related legal systems. Given that people living in Japan can easily access and use AI services which overseas business operators provide via the Internet, on a daily basis, and that domestic business operators may suffer unilateral disadvantages if legal systems which impose business operators any obligation exclude only overseas business operators, overseas business operators should also be covered by the legal system same as domestic business operators. In this case, to ensure the effectiveness of the legal system even for overseas business operators from whom it is difficult to obtain compliance cooperation due to geographical factors, etc., consideration should be given to formulating rules that clearly include overseas business operators as well. In practice, if an overseas business operator has a branch office or representative in Japan, it may be possible to consider requesting responses through them.

3．Promotion of innovation while mitigating risks

While AI can cause various risks depending on the methods of its development or use, etc., it also has the potential to greatly contribute to improving people's lives and developing the national economy. It is important to promote innovation while mitigating risks to make Japan the friendliest country in the world for developing and implementing AI.

（1）Promoting innovation

① Support for research and development

Research and development of AI require large amounts of training data, as well as facilities and equipment that can handle large-scale information processing, information communications, data storage, etc. From the perspective of aiming to promote innovation by various stakeholders

(including startup companies) using these facilities and equipment, it is important that the government makes progress in the promotion. In addition, with regard to human resources researching and developing, since the shortage of AI talent in Japan is becoming more serious due to the intensifying global competition for talent, it is important to actively promote human resource development.

The government and other relevant stakeholders are already working on organizing and expanding high-quality Japanese language data, etc. and providing it appropriately to Japanese companies. They are also working on promoting the establishment of data centers, securing power supplies, and supporting AI semiconductors. Human resource development projects such as" Next Generation AI Human Resources Development Program[3]" have been implemented. Support for research and development of AI should continue to be carried out from various aspects.

In addition to these, as seen in the development of generative AI and rapid expansion of its usage following breakthroughs in basic research and development related to large-scale language models, there is a close relationship between basic research of AI and its broader application. Therefore, it is necessary to consider the promotion of basic research.

At present, RIKEN is a National Research and Development Agency working on AI for Science, which involves using AI foundation models in scientific research, and this has the potential to bring about major changes in scientific research methods and research itself. Further, in April 2024, Japan-U.S. Joint Leaders' Statement involved collaboration between Argonne National Laboratory, U.S. and AI for Science, and international collaboration has also been carried out. It is important to continue to engage in such research and development activities that contribute to the development of other fields.

②Use by business operators

To improve the competitiveness of Japan, it is important to widely implement AI technology in society and ensure its utilization by domestic business operators as well as to support research and development as referred to in ①. For example, promoting initiatives that transfer and apply research results from government research institutions, universities, etc. to society can lead to new business stakeholders by business operators and revitalization of various industries.

In Japan, it is common for business operators to develop their own AI systems and provide AI services for consumers. As it can be considered that the number of business operators that

[3] "Broadening Opportunities for Outstanding Young Researchers and Doctoral Students in Strategic Areas; Next Generation AI Human Resources Development Program" is a program run by the Japan Science and Technology Agency that provides support to young researchers and doctoral students who are working on research and development in the field of next-generation AI: AI fields and emerging and integrating domains in the AI fields.

provide AI services for consumers by using AI systems developed by other companies[4] will much increase in the future, it is also important for the government to create an environment that allows such business operators to smoothly provide AI services both domestically and internationally. Specifically, to promote highly secure AI systems to expand their market, it is effective to ensure international consistency and implement safety evaluations and certifications, etc. as mentioned later. For example, by providing incentives for those who have obtained certification and increasing the number of the certified business operators, it is considered that many citizens will be able to use AI more securely, the market will expand, and innovation will be promoted. Furthermore, to encourage domestic business operators to participate newly in the AI market, it is also important to create an environment where business operators can learn the basic knowledge, etc. of AI.

In addition, we should also focus on use of AI in robotics, medicines and disaster prevention, and international collaboration and contribution such as collaboration with Asian countries.

（2）Application of laws and making use of soft law

Main laws and regulations related to examples of risks that AI can pose, are as shown in Figure 3. There are some real cases occurred in Japan: the face of a performer who was in a sexual video content was replaced to that of the different entertainer and was disclosed in the Internet (violation of defamation in the Penal Code, and of the Copyright Act), computer virus was created by utilizing generative AI (the Crimes Related to Electronic or Magnetic Records Containing Unauthorized Commands in the Penal Code), and a woman character created by utilizing generative AI was misused to perpetrate money fraud (the Crimes of Fraud in the Penal Code). On the premise of the fact that such risks have already been mitigated to some extent by existing laws, it is necessary to consider further about legal systems.

In mitigating risks, there are responses by laws and by soft law such as guidelines.

Currently In Japan, to deal with risks and problems caused by AI, the government ministries and agencies responsible for each field are responding through laws and soft laws. For example, in June 2023, while indicating points of caution related to handling personal information when using generative AI services, the Personal Information Protection Commission alerted personal information handling business operators and administrative bodies[5] to handle personal

---

[4] Benesse Corporation is using AI systems developed by other companies to provide services for customers, such as “AI to help with Independent Research” and services that use accumulated knowledge and data (as referred to Material 2 at the 2nd meeting of the AI Institutional Study Group).

[5] This refers to administrative bodies, local governments’ bodies (except councils), incorporated administrative agencies (except agencies listed in Appended Table 2, the Act on the Protection of Personal Information) and local independent administrative agencies (except institutions whose main purpose is the

information properly, in accordance with the Act on the Protection of Personal Information (Act No.57, 2003). Further, in March 2024, the Legal Subcommittee under the Copyright Subdivision of the Cultural Council, Agency for Cultural Affairs, presented a certain way of thinking related to the interpretation of the relationship between AI and the current Copyright Act in mitigating concerns about potential copyright infringement when generative AI learns from a great amount of data and generates contents. Additionally, in May 2024, the Intellectual Property Study Group, the Cabinet Office, summarized the thoughts on legal rules on the relationship between AI and intellectual property rights, while indicating the thoughts on the necessity that each actor works on through a combination of legal, technical, and contractual measures to realize an ecosystem where the advancement of AI technology and the appropriate protection of intellectual property rights can coexist. Furthermore, as mentioned at the beginning of this report, in April 2024, the Ministry of Internal Affairs and Communications and the Ministry of Economy, Trade and Industry published the “AI Guidelines for Business Ver1.0”[6], which states that AI systems and services should be developed, provided, and used respecting the rule of law, human rights, democracy, diversity, and a fair and just society.

If there is any penalty provided in laws, public organizations can invoke some kind of coercive power, which means an advantage that it is easy to ensure effectiveness of the rules, but, it may hinder the development of the regulated field, and there is also the drawback that it lacks flexibility, as it takes a certain amount of time to consider their scope by the fact that the regulations which affect the rights and interests of citizens need to be clear. Even laws that do not involve penalty can discipline domestic and overseas business operators and thereby ensure a certain level of effectiveness by clearly stating their obligations and responsibilities in the laws.

On the other hand, soft laws such as guidelines has the advantage of being able to respond quickly and flexibly in line with the international situations and the latest technological trends, and having less negative impact on innovation, while the soft law might have to rely on the voluntary response of business operators.

In general, Japanese companies are known for their high consciousness of legal compliance, and when new regulations are established, there is a possibility that they may hesitate more than necessary in conducting new research and development or in developing and deploying new services due to their focus on adhering to these regulations. In the field of AI, where technological development and service changes are rapid, excessive regulations that inhibit research and development, or development and deployment of services carry the risk of undermining Japan's international competitiveness in the future. Therefore, it is necessary to fully consider the impact on innovation when considering and introducing new legal systems.

---

business listed in Article 21, item 1 of the Local Independent Administrative Agencies Law, or whose purpose is the business listed in the same Article, item 2 or 3(h)).

[6] In November 2024, the version 1.01 was released with updates.

From what is mentioned above, it is essential to appropriately combine laws and soft laws such as guidelines to ensure the promotion of innovation while mitigating risks. Basically, the autonomy of business operators should be respected, and regulations by laws should be limited to cases where voluntary efforts by business operators cannot be expected.

Furthermore, in domains where individual laws and regulations already exist, AI is beginning to be used for various purposes in each domain, and the circumstances where rights and interests need to be protected vary depending on the application of AI, so it is essential to first address these issues by making use of the framework of the relevant laws and regulations. In addition, if implementing legal regulations including the cases of creating new legal systems outside of such domains mentioned above, while taking into account the impact on the activities of business operators and the risks posed by AI, as mentioned in (3) below, the extent of application should be at the level which protects the rights and interests that truly need to be protected. In doing so, it is important to be aware of the division of roles between the government and business operators, and to clearly draw a line between what is subject to regulations and to what extent business activities are permitted. It is also important to consider based on the principle of technological neutrality in regulations, which means that regulations should not mandate or prioritize the use of specific types of technology in order to achieve their objectives. Applying regulations to cases where inappropriate AI is prototyped for legitimate research related to AI safety, etc., needs to be considered including its necessity. When considering legal systems that broadly target business operators in general, it is necessary to ensure that the systems are manageable for business operators of any scale, including startup companies, by considering the burden on business operators associated with compliance.

Examples of risks that AI can posec	Specific examples and hypothetical cases	Main laws, etc.
Entering confidential information into AI	Entering confidential corporate information into an external AI service, ending up with information leakage	The Unfair Competition Prevention Act, The Civil Code（contract） ※February 2024, the points of notes when using AI were summarized in "The Handbook to protect confidential information" (Ministry of Economy, Trade and Industry)
Infringement of others' copyright in the process of developing, training, generating and using AI	Training for the purpose of generating images similar to illustrations of specific manga or, anime characters, etc., and generating and using such images similar to those illustrations	The Copyright Act ※March 2024, “General Understanding on AI and Copyright in Japan" was published, clearing the interpretation（Agency for Cultural Affairs）
Infringement of others' industrial property rights in the process of developing and using AI	By training the registered trademarks of others, creating trademarks that are identical or similar to the registered trademarks, and using them for goods or services that are identical or similar to the designated goods or services	The Design Act, The Trademark Act ※May 2024, "Intellectual Property Right Study Group in the AI era Interim Report" was published, organizing the approach to legal rules.（Intellectual Property Right Study Group in the AI era）
Privacy violations and violations of personal information protection in the process of developing and using AI	Using data that includes personal information for AI training without the consent of the individual	The Constitution（privacy right, publicity right), The Act on the Protection of Personal Information ※June 2024, "the Act on the Protection of Personal Information the Every-Three-Year Review Interim Report" was published, organizing points of discussion in using AI（Personal Information Protection Commission）
Malfunction of products with AI	Safety of lives and bodies are impacted by malfunction of automatic driving vehicles	The Road Transport Vehicle Act, The Act on Pharmaceuticals and Medical Devices (PMD Act), The Industrial Safety and Health Act, The Civil Code (torts, etc.), The Product Liability Act, The Act on Securing Compensation for Automobile Accidents, The State Redress Act
Deepfake (misuse of portraits, voices, etc. synthesized by AI)	The act of synthesizing and disseminating personal images into pornographic or other sexual images without the person's consent, and fraud through voice calls pretending to be a celebrity or acquaintance using AI	The Civil Code (moral right, torts), The Penal Code(crimes of intimidation, defamation, distribution of obscene objects, fraud, damage to credibility; obstruction of business, etc.), The Child Pornography Prohibition Act, The Information Distribution Platform Act (information on infringement of rights)
Promotion of bias (discrimination and prejudice)	Implementing human resources assessments and judgements related to hiring and retiring by inappropriate AI	The Act on elimination of hate speech, Labor-related Laws and Regulations, The Civil Code, The Personal Information Protection Laws, The Act for Eliminating Discrimination against Persons with Disabilities, The Act on the Promotion of the Elimination of Buraku Discrimination
Information manipulating by disinformation and misinformation	Disturbing election by creating false information about the candidates using AI and spreading it on SNS, etc.	The Civil Code (moral right, torts), The Penal Code（crime of defamation), Administrative Regulations、Public Offices Election Act, The Information Distribution Platform Act (information on infringement of rights)
Violation of the rights and interests of citizens	There is a possibility that individuals may suffer disadvantages, such as being unable to receive administrative services, due to incorrect judgments made by AI	The Constitution (due process), The Administrative Procedure Act
Cyber attacks such as creating viruses	Creating computer viruses by misusing generative AI	The Penal Code (crimes related to electronic or magnetic records containing unauthorized commands), The Act on Prohibition of Unauthorized Computer Access
Hallucination (AI creating false information)	Generative AI creating false information to mislead users	The Civil Code (torts, contract)
Increased environmental impact	Increasing CO2 emissions due to increased power demand, etc. in the process of AI development	The Act on Promotion of Global Warming Countermeasures
Negative interaction between humans and AI	A person who became engrossed in conversations with AI became pessimistic about life and commit suicide.	The Basic Act on Suicide Countermeasures
Concerns that AGI might go out of control	There is a possibility that humans might lose control of AGI, causing social chaos	-

※"The AI Guidelines for Business Ver1.01" describes several of the above risks and indicates how to respond to them under 10 common guidelines (human-centered, safety, fairness, ensuring privacy, ensuring security, transparency, accountability, education and literacy, ensuring fair competition, and promoting innovation).

Figure 3 Categorizations related to examples of risks that AI can pose.

（３）Mitigating risks

When mitigating risks, it is effective to establish separate standards in using AI in specific domains as well as to clarify common contents with which AI stakeholders should comply.

At present, the OECD "Recommendation of the Council on Artificial Intelligence" (the OECD AI Principle) and Hiroshima Process International Guiding Principles, etc. exist as the international framework. Based on these, as a domestic norm, "AI Guidelines for Business" for all those involved in the development, provision and use of AI across various business activities, was published, and what should broadly and generally be complied with has been formulated. In order to mitigate risks more appropriately as well as to ensure the promotion of innovation, while updating the content of the “AI Guidelines for Business” in line with technological developments, etc., it is important to deal with as necessary as well as to promote dissemination and enlightenment, etc. so that each entity complies appropriately. To appropriately mitigate risks, in addition to clarifying the roles of each entity such as developers and providers in accordance with the guidelines, it is necessary to clarify responsibilities. At the same time, it is necessary to share vital information and closely collaborate among respective stakeholders such as between developers and providers, and between providers and users.

When using AI in a specific domain, it is important to consider the purpose, method of use, etc., and to respond individually. For example, AI related to infrastructure and services that form the basis of people's lives and economic activities (hereinafter referred to as “infrastructure services, etc.”), and product safety is addressed primarily through existing laws by the respective ministries and agencies.

In addition, due to the rapid development of AI, it is also necessary to appropriately mitigate newly emerging risks depending on the content of respective fields. In particular, as for AI that actually poses or is highly likely to pose a serious threat to fundamental human rights and interests such as life, body, and property, as well as to social safety and national security, the necessity of regulation should be considered depending on the contents of the risks and the severity of their social impacts.

To achieve this, collaborating between the government and business operators, based on the actual cases, with various AI models and uses in existence, at each stage of the AI lifecycle, such as development, provision, and use, it is necessary to analyze the factors of potential risks, including the types of AI models, the nature of risks they pose, and the impact they may have on different stakeholders. As a prerequisite, it is important to survey and analyze the current state of AI development and use, and then to take necessary measures timely and appropriately after sharing this awareness throughout society. For example, there are concerns about inappropriate AI screening job applicants and AI confusing consumers, and the government should strive to consider necessary measures including technical responses while working on comprehending the actual situation. Further, when requesting cooperation from each entity for the government’s understanding of the actual situation, the government should adhere to the basic principles such

as the rule of law, due process, democratic and responsible administrative, and fundamental human rights, especially the right of privacy etc., as confirmed in the Hiroshima AI Process and so on, and it is necessary to prevent the government from arbitrarily exercising its authority and reducing predictability or causing a shrinkage effect on business operators.

Other countries have established regulations based on the scale of AI including the amount of calculation of training, or the number of users, but in light of the development of high-performance AI that is not dependent on its scale, it is necessary to consider what factors should be taken into account.

４．Promoting international cooperation

（１）Formation of AI governance

With regard to AI governance, there are lively discussions in framework of multilateral countries.

As mentioned in “1. Preface,” following the G7 Hiroshima Summit in May 2023, G7 established the “Hiroshima AI Process” to consider international rules regarding generative AI. In December of the same year, the "Hiroshima AI Process Comprehensive Policy Framework", which includes the “Hiroshima Process International Guiding Principles for All AI Actors” and the “Hiroshima Process International Code of Conduct for Organizations Developing Advanced AI Systems”, was approved by the G7 leaders. After that, Japan has been sharing the spirit of the Hiroshima AI Process at various international conferences, etc., aiming at realizing safe, secure, and trustworthy AI. And as an outreach beyond G7, in May 2024, with the support of 49 countries and regions that agree with the spirit of the initiative, the “Hiroshima AI Process Friends Group” was established, aiming to expand support. As the international formation of AI governance will determine the direction of future AI development and contribute to national interests, Japan should continue to lead the discussion at various international conferences based on the concept of Hiroshima AI Process. In addition, Japan should establish an AI legal system that serves as a model for other countries and disseminate it to the world.

In September 2024, the “Global Digital Compact” was adopted as an annex to the outcome document of the Summit of the Future at the United Nations. It included the establishment of an international scientific panel on AI to promote scientific understanding through the assessment of AI risks and opportunities, as well as the launch of a global dialogue on AI governance that national governments and related stakeholders participate. Additionally, the Council of Europe has established the “The Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law” and, at present, in addition to the EU, 10 countries including the United States and the United Kingdom have signed the convention. The OECD established AI principle in 2019, which consists of inclusive growth, sustainable development and well-being, human-centered values and fairness, etc., and has just updated it

in May 2024 to mitigate risks of disinformation and misinformation by generative AI developing rapidly. The Global Partnership on Artificial Intelligence (GPAI), which is an international and multistakeholder initiative to realize the development and use of “responsible AI” based on a human-centered approach, in November 2022 held the GPAI Summit 2022 in Tokyo and reached an agreement of the first Ministerial Declaration among respective countries on the promotion of the use of AI based on human-centered values, opposition against the illegal and irresponsible use of AI, and contributions to sustainable, resilient, and peaceful societies, etc., whose result was announced as the first ministerial declaration. After making an integrated partnership with the OECD in July 2024, GPAI members have reached 44, and the first GPAI Expert Support Center in Asia was established in Tokyo to support projects regarding safety evaluation of generative AI promoted by the Hiroshima AI Process. Furthermore, various discussions are taking place internationally, and in terms of implementing legal systems and policies related to AI, we should respond based on agreements or acceptable arrangements and concepts in these international frameworks, etc.

（2）Ensuring international consistency and interoperability

As mentioned earlier, citizens can use AI services provided by business operators in various countries through the Internet, and AI services in Japan can be used by people from all over the world. In this situation, if the interoperability between international norms regarding safety, etc. that must be satisfied and norms applied in Japan is ensured, business operators in Japan can smoothly advance into overseas markets, and citizens are able to access AI services all over the world, so it is important to ensure international consistency and interoperability.

Therefore, in addition to the importance of the Hiroshima AI Process, standardization activities that set international standards in ISO, IEC, etc. are significant for the development of global markets and business activities in the future, and it is necessary to discuss among a wide range of stakeholders, have a long-term perspective and promote actively.

The AI Safety Institute (AISI), an organization that examines and promotes evaluation methods and standards for AI safety, is working as a hub for AI safety knowledge in domestic and overseas on building a network with related domestic and overseas organization, as well as creating guidance for safety evaluation, etc. Following the AI Safety Summit held in the UK in November 2023, discussions on governance from a technical perspective regarding the safety of AI progressed, and in February 2024, the AISI was established in Japan following the UK and the US, and to ensure the international consistency and interoperability mentioned above, it is important to promote initiatives by the AISI.

III. Direction of specific legal systems and policies

1. Overall items

AI, including generative AI, is highly versatile and is used in various fields, and there are different ways to mitigate risks. With regard to mitigating risks including handling of personal information or copyright and dealing with disinformation and misinformation, it is on the premise of responding mainly through existing laws, etc.. In terms of AI, however, there are also cases where a cross-cutting response is required, so it is necessary to strengthen the government's function of a strategic leadership board which oversees the whole picture, to formulate strategies, and to ensure the transparency and appropriateness to improve safety, and, as necessary, it is appropriate to develop legal systems.

（1）Strengthening the government’s function of a strategic leadership board and formulating strategy

Due to the expansion of fields of application and uses, the emergence of general-purpose AI, etc., there are cases where the period from research and development to use is short, and initiatives at each stage during this period can be carried out almost simultaneously. For this reason, the initiatives in various stakeholders and processes from research and development to use are closely relating each other, so it is necessary to carry them out in integrated and cross-cutting manner, and the government's function of a strategic leadership board should be strengthened to promote integrated measures from research and development to use in economical society.

In addition to the case of using AI to improve people' lives in various situations including by the national government and local governments, there are also concerns over its misuse for criminal purposes, and it also has an aspect of dual-use technology. So, in terms of strengthening the function of a strategic leadership board, it is necessary to establish a system of policy promotion in which the relevant government ministries and agencies participate widely.

Furthermore, to promote comprehensive measures, it is necessary for the strategic leadership board to formulate a strategy or a basic plan (hereinafter referred to as “the strategy”). As for AI, ensuring safety and security is important for promoting the use of AI, promoting innovation, responding to national security risks, preventing crimes, etc., so the strategy should contribute to promoting safe and secure research and development and use of AI. The strategy should include measures that require the whole government to work together to promote innovation with mitigating risks while ensuring international cooperation.

The above should be legislated to strengthen the function of a strategic leadership board for AI and clarify its authority that the board can request cooperation from related administrative agencies.

（２）Improving safety

To improve the safety of AI, it is necessary at least to ensure the transparency and appropriateness in the lifecycle from research and development to use. It is also considered that it will be effective to utilize a third-party certification or safety evaluations which business operators voluntary undertake. Furthermore, the government should survey the actual status regarding a technology of AI rapidly promoting and its usage trends to provide information, and, as necessary, request relevant stakeholders to take actions.

To implement these measures, it is necessary for the public and private to cooperate and work together because information sharing and cooperation from related stakeholders including business operators are essential.

① Ensuring the transparency and appropriateness through AI life cycle

From research and development to use of AI, for example, in the case of using a large amount of training data to construct models, performing fine-tuning, transferring them from developers to providers, and then trained further by providers, and the AI services are provided to users, the ability of the AI systems which developers developed and that of the AI systems when providers provide the AI services to users are likely to be different and the result of risk evaluation also might be different. While users cannot always grasp how to mitigate risks at the point of development, when mitigating risks, if the necessary information regarding such risks is not appropriately shared with stakeholders, there is a possibility that providers will provide AI systems to users based on misunderstanding, or that users will use AI services inappropriately, resulting in the risks becoming apparent.

For this reason, to ensure the safe and secure research and development and use of AI, transparency should be ensured to share the necessary information between developers and providers, and providers and users. On the other hand, at this stage of research and development, it has not been yet put to a practical use, and information regarding research and development is likely to be related to confidential information of companies. It is thus important to limit information sharing to a truly essential extent so that measures to ensure such transparency try not to fall into excessive burdens on business operations or extensive information disclosures of business operators.

As for appropriateness, the “Hiroshima Process International Guiding Principles for All AI Actors” agreed upon in the Hiroshima AI Process required AI actors to identify and mitigate risks and incidents when developing advanced AI systems and before and after putting them into the market, and to promote the trustworthy and responsible use, etc., and various discussions are being held in AISI, ISO, etc. in other countries, and it is necessary to promote appropriate research and development and use.

To ensure appropriateness, it is appropriate for the government to develop guidance based on the spirit of international norms such as the Hiroshima AI Process, and to encourage business

operators to take voluntary actions in accordance with various norms. As for the guidance, it is important to consider that, for example, AI developers have to construct measures to constrain inappropriate output, and disclose and share information related to concept or risks of AI and proper transparency of training data sets should be supported.

Furthermore, to ensure transparency and appropriateness, the government should ascertain the situation of business operators by surveys and should provide necessary support including responses in accordance with existing laws based on the results. As it is not possible for the government to ascertain the situation of business operators or provide necessary support without the cooperation of business operators, it is appropriate to respond by legal systems so that the government can request a cooperation including information sharing from domestic and overseas business operators.

As countermeasures against disinformation and misinformation, there are some examples in other countries where it becomes a part of obligations that customers are informed that services are with using AI or that contents created by AI are indicated as a creation by AI, and it needs to be paid attention to. In addition, to ensure transparency and appropriateness, technical responses such as digital watermarking[7] or digital provenance[8], etc., are also important. In relation to this, in September 2024, the Ministry of Internal Affairs and Communications "Study Group on How to Ensure the Soundness of Information Distribution in the Digital Space," pointed out that with regard to disinformation and misinformation, it is important to promote research and development, and social implementation of technology that judges whether information is generated by AI or not, or technology ensuring trustworthiness of information contents and their senders, and the study group proposed specific measures to the government.

② Strategic promotion related to safety evaluation and certification implemented by domestic and overseas organizations

In promoting the safe and secure use of AI, the implementation of safety evaluation and certification systems is one of effective means. Safety evaluation and certification systems can be generally categorized into those related to the evaluation and certification of AI systems and those related to the evaluation and certification of governance, etc. for organizations using AI.

As for the safety evaluation of AI systems, developers, providers and users of AI who understand the risks should basically carry out their own risk assessments and take actions by utilizing internal and external expert teams or evaluation tools. If a useful third-party certification is established in the future, it is conceivable that AI developers and providers will obtain such

---

[7] Digital watermarking is a technology that embeds identification information in contents to show that it is generated by AI, such as “SynthID.”

[8] Digital provenance is a technology that attaches provenance information of the creators, etc. to the contents in a verifiable form, such as “Originator Profile” and “C2PA.”

certification, and that many users, including general citizens, and providers who have not handled AI services so far, will also evaluate the safety of AI. Providers and users will be able to recognize and select AI business operators with high levels of security or AI systems which these operators create, based on whether or not they are labeled by the third-party certification.

On the other hand, with regard to governance for organizations using AI, there will be cases where the users build system by themselves and evaluate on their own, or a third-party certification is used. A third-party certification system is thought to be one of the measures that will promote the safe and secure use of AI and contribute to the vitalization of AI industry in Japan.

Governance certification of AI systems and organizations using AI is currently being considered in the context of ISO.

Safety is already regulated by individual laws such as the Electrical Appliances and Materials Safety Act (Act No. 234, 1961) and the Act on Securing Quality, Efficacy and Safety of Products Including Pharmaceuticals and Medical Devices (Act No. 145, 1960), and also responded by "AI Guidelines for Business". Furthermore, discussions are also being held in AISI communities of each country, ISO, IEC, and other organizations.

Activities to create international criteria and standards are important for ensuring the interoperability of AI systems in the future global market and need to be addressed proactively and strategically.

Furthermore, the development of legal systems in Japan should be based on international norms and taken into account the effectiveness of the systems. When implementing AI evaluation and certification, if their levels can be set according to the users and purposes of use, or if a system which reduces the burden on users to confirm a certain level of safety can be established, or a system which certifies organizations that implement the evaluation and certification can be established, it is conceivable that the system will become more effective and sustainable. However, when building this system, on the premise of the activities of AISI, ISO, and other initiatives, it is necessary to consider in detail which stakeholders are involved and what criteria are used for evaluation. In order to ensure AI safety, AISI is expected to continue to work with relevant ministries on research, analysis, organization, and information dissemination, and to support organizations that will serve as a strategic leadership board.

③ Investigation and information dissemination by the government regarding serious incidents.

As mentioned above, AI has been developing rapidly in recent years, and various risks are increasing. In this situation, in order to mitigate the risks of AI and implement appropriate measures as the government, it is appropriate for the government to provide information to citizens to essential extent so that first, the government collects and ascertains information of

actual status related to development, provision and usages of AI which has been changing from aspects of both technology and business activities. And then business operators can use AI effectively and appropriately while citizens can deepen their understanding and being interested in promotion of research and development, and use of AI, at the same time when companies fulfill accountability while the government respects their confidential. In particular, with regard to AI models that many citizens use daily, the government should collect information related to safety and the transparency of AI including risks measures of supply chain and provide the information broadly to citizens, and then users will be able to recognize and select AI providers with high levels of safety and their AI systems. Besides, it is important for the government to collect information regarding actual status of AI introduction in infrastructure services, etc.

In addition, if serious accidents caused by use of AI actually occur, the government will need to take measures to prevent its escalation or recurrence, as well as to raise awareness of recurrence prevention measures by business operators that develop and provide AI. In other words, in terms of AI used in Japan, in the event where it is detected that serious problems which infringe on the rights and interests of citizens arise or are likely to arise, the cause of problems should be investigated, and, as necessary, guidance and advice should be provided to stakeholders, and the investigated information should be made known to citizens. It is important to judge whether or not it can be said that incidents have occurred, based on cases or knowledge piled in the government through the information collection and ascertainment mentioned above.

This survey and dissemination of information cannot be carried out without the cooperation of business operators, so it is appropriate to respond by legal systems so that the government can require domestic and overseas business operators to cooperate to provide information.

2. Use by the government

The rates of using AI by individuals and companies in Japan are significantly lower compared to in other countries[9]. AI is expected as a foundation of the development of people's lives and economic activities to increase importance of its use. However, if the situation is overlooked, it is likely to undermine Japan's international competitiveness instead. For this reason, it is considered that, to begin with, the government should take the lead in using AI and then encourage citizens to use it.

[9] According to the 2024 White Paper on Information and Communications by the Ministry of Internal Affairs and Communications, in Japan, only 9.1% of individuals use generative AI (China (56.3%), the U.S. (46.3%), UK (39.8%), and Germany (34.6%)), and in a survey to companies, the percentage of companies using generative AI in their business was 46.8% (the U.S. (84.7%), China (84.4%), and Germany (72.7%))

（1）Government procurement

The government's indication of basic approach to using AI is also useful from the perspective of encouraging AI developers and providers who are considering participating in government procurement to take their own initiatives to enhance safety.

When the government procures information systems, based on the “Agreement on the Procurement Policy and Procurement Procedures for Goods, and Services of the State related to IT Procurement” (agreement between relevant government agencies on December 10, 2018), regarding procurement deemed necessary to respond to supply chain risks, actions are to be taken by necessary measures after consulting with National center of Incident readiness and Strategy for Cybersecurity and Digital Agency, and AI is also subject to the agreement.

On the other hand, so far, there are no guidelines that are specific to AI procurement, but there are various risks associated with AI, as mentioned above. Therefore, to mitigate the risks, it is important to develop government procurement guidelines specific to AI that can be used as a reference when the government procures AI, or to deepen operation of the existing guidelines related to AI.

In addition, the development of guidelines related to such government procurement may be able to contribute business operators using AI to utilize some kind of AI systems or services, leading to promotion of the spread of AI with high levels of safety.

In developing guidelines, while aiming to mitigate risks and assure quality of AI, if ensuring transparency is required, companies' burdens should be taken into consideration according to use cases. This should be done in a way that allows for easy and quick participation by many business operators.

（2）Use by the government

The government can improve the quality and efficiency of administrative services and operations by using AI. In addition, the government can promote AI use and can contribute to the development of the domestic AI market by showing use cases and AI usefulness, and informing specific examples and points to note. Therefore, it is important that the government takes the lead in using AI. However, matters that could have a significant impact on the rights and interests of citizens should be carefully considered in light of the risks[10] of automatically adopting the output results of AI.

---

[10] There are some cases of risk occurrence overseas: the case of unemployment insurance in the United States was that the integrated data automation system used by the Michigan Department of Unemployment Insurance to detect fraud by claimants recorded 93% of errors over a two-year period from 2013, falsely accusing 20,000 state residents of fraud, and the case of teachers losing their jobs was that the Houston Independent School District introduced a value-added model (VAM) to estimate the impact of

Local governments also account for a large part of administrative services, and they have a significant impact on people's lives, so it is important to promote the use of AI and enhance their quality and efficiency of administrative services and operations, etc. In addition, when local governments promote the use of AI, they can refer to examples of use including the advanced initiatives of each local government[11], and the use of AI in response to the regional issues of each local government is also significant.

As for the use of AI services by the government, etc., according to “Agreement on the Business Use of Generative AI such as ChatGPT (2nd Edition)” (September 15, 2023, Agreement of the Executive Board of the Digital Society Promotion Council), the government cannot handle confidential information on cloud services that are provided to unspecified large number of users and are available only by agreements to standard contracts or rules. In cases of considering the use of generative AI based on individual contracts, etc., the government can handle some of the class 2 confidential information on cloud services only if they conduct to analyze risks properly after obtaining approval from the government's AI Strategic Team. When the government uses AI, it needs to pay attention to handle confidential information and follows the agreements mentioned above, and as necessary, it is important to update by referring to initiatives[12] in other countries.

Furthermore, it is significant to implement measures for reskilling or literacy improvement to encourage citizens to use AI.

3. Matters related to life and body safety, systemic risks and national security

It is considered necessary to particularly focus on dealing with matters such as medical devices, self-driving cars, and infrastructure services, etc., that are related to people's life and body safety as

---

teachers on students' academic growth, and used it to deny contract update or dismiss teachers from 2007 to 2016, so many teachers sought judicial relief to stop the use of VAM.

[11] Kobe City established the AI ordinance in March 2024, aiming at effective and safe use of AI in accordance with certain rules, and the city has also been promoting the use of AI such as summarizing text, formulating ideas, and generating programming codes (as referred to Material 3 at the 2nd meeting of the AI Institutional Study Group).

Also, Tokyo Metropolitan Government has been promoting initiatives: it established and revised the guideline and held Tokyo Metropolitan Government AI Strategic Council in December 2024.

[12] Other countries provide guidelines and training programs of the use of AI for their government agencies . For example, the Canadian government published the guide of using generative AI for federal government officers in September 2023. In addition, the government of California, the U.S., published the guideline in March 2024 for generative AI products that state government agencies purchase, and began providing training programs on the procurement of generative AI for state government officers. The EU AI Act, which passed in May 2024, also includes the evaluation and promotion of best practices in public procurement procedures related to AI systems as one of the “measures that the AI Office takes.”

well as systemic risks[13] which specifically have a significant impact on people's lives and social activities, but each government ministry and agency responsible for each sector responds based on existing business acts, and as necessary, they are also in a state of dialogue with the business industries regarding the development or use cases of AI technology to determine whether or not additional measures are needed.

So far, each ministry and agency should continue to respond under the existing laws and guidelines, but in the future, if new risks emerge and cannot be dealt with by using existing frameworks, the government should consider reviewing legal systems or establishing new legal systems after clarifying the interpretation of relevant frameworks. As for systemic risks, large-scale AI systems in which multiple AI systems work together may support social systems in the future, and since if such groups of AI systems behave in unexpected ways, it could cause great confusion throughout the society, it is important to deal with such cases appropriately.

Furthermore, in terms of risks regarding national security such as the development of CBRN weapons or the use of AI for cyberattacks, it is necessary for the relevant ministries to further consider the necessary responses from the perspective of ensuring Japan’s national security.

IV. Conclusion

While legal systems related to AI are being developed in other countries, considering the above, Japan should establish guidance for widely used AI to achieve the promotion of innovation while mitigating risks. Additionally, while encouraging business operators to take the lead in ensuring the transparency and appropriateness, on the natural premise of ensuring the safety of life and body, and national security, the government should investigate and ascertain conditions of various risk mitigations, including actual measures taken in development and use of AI. Moreover, if serious problems occur or are likely to occur, the government should take actions based on the existing laws, provide necessary support, or other actions. Specifically, technological innovation of AI has been occurring rapidly around the world, and therefore, the government needs to conduct systems which allows it to ascertain and respond situations as soon as possible. As for the procurement and use by the government as well as infrastructure services, etc., while continuing dialogue with business sectors, the government should firstly respond through relevant acts and guidelines. Like this, it is important to establish risk governance through the cooperation of public and private sectors.

In establishing and handling the above-mentioned guidance by the government, and investigating and comprehending actual status related to AI, while voluntary actions by business operators are important, it is necessary to ensure effectiveness. Therefore, considering the impacts on business operators' activities, the implementation should be carried out through legal systems. In introducing

[13] They are the risks that malfunction in specific systems will spread to relevant systems, causing serious impacts on wide areas.

these legal systems, it is necessary to comply with the basic principles of rule of law, due procedure, democratic and responsible administrative, and fundamental human rights, especially the right of privacy etc., which confirmed in the Hiroshima AI process, and to be careful of ensuring that legal systems do not hinder promoting innovation of AI.

Based on this report, we expect the government to promptly implement social systems, including legal systems related to AI, to make development and usage of AI the friendliest, to be based on human-centered values which serve as a model for other countries.