

International Collaboration in AI Governance Key Considerations of the Council of Europe's AI Convention and Japan's Response

Arisa Ema

Associate Professor, Institute for Future Initiatives, The University of Tokyo
Visiting Researcher, RIKEN Center for Advanced Intelligence Project

Tagui Ichikawa

Specially Appointed Professor, Institute of Innovation Research, Hitotsubashi University

Toshiya Jitsuzumi

Professor, Chuo University / Visiting Researcher, Georgetown University

Hiroshi Nakagawa

Team Leader, RIKEN Center for Advanced Intelligence Project

Hiroki Habuka

Research Professor, Graduate School of Law, Kyoto University / CEO, Smart Governance Inc.

Yuko Harayama

Professor Emeritus, Tohoku University
Visiting Researcher, Institute for Future Initiatives, The University of Tokyo

Executive Summary

Currently, artificial intelligence (AI) technologies are widely deployed with high expectations of transformative impacts on daily life and work. Consequently, owing to its significant impact it may have on society, AI governance has become a prominent subject of discussion in international organisations and networks, including the G7 Hiroshima Summit. The Hiroshima Summit mandated the creation of the G7 Hiroshima AI Process to harmonise various regulatory frameworks and promote AI governance in a coordinated manner, particularly with regard to generative AI. In addition, the Ministerial Declaration of the G7 Digital and Tech Ministers' Meeting acknowledges the need for agile and decentralized governance with multi-stakeholder participation, and legal frameworks, designed to operationalise the principles of the rule of law, due process, democracy, and respect for human rights.

The Council of Europe (CoE) is engaging in discussions concerning the Convention on AI, which considers values such as the rule of law, human rights, and democracy in the life cycle of AI, from design to decommission. Negotiations are currently underway to draft the world's first convention on AI, which includes non-member states participating as observer states. A framework convention, as such, while possessing legally binding authority, grants contracting parties autonomy to determine the means to achieve convention objectives. Key provisions include taking measures to ensure the availability of effective remedies for violations of human rights and fundamental freedoms, record-keeping for such incidents, and the implementation of risk and impact assessments based on a risk-based approach. The Convention also stipulates the establishment of oversight mechanisms responsible for monitoring compliance.

All G7 countries participated in the drafting and negotiation of the AI Convention, and Japan, as an observer state, is in a position to influence the formation of AI-related rules. Considering this, Japan, as the chair of the G7 Hiroshima AI Process, should position the Convention as a tool for expanding AI-related rules worldwide based on the shared values of the G7 countries and be actively involved in negotiations. Given this context, this policy recommendation identifies the points that Japan should clarify its positioning in view of the negotiation process toward the conclusion of this Convention and proposes the following three points.

AI-related regulations that would be included in the scope of this Convention should be identified, and proactive participation in the negotiation process is essential

Many provisions make it difficult to determine whether new legislation is required to align with this Convention or whether existing laws, regulations, and guidelines can be addressed without new legislation. In Japan, in particular, horizontal regulations on the development and use of AI systems have been addressed by guidelines, rather than legislation, and existing domestic laws pertaining to AI are limited. In addition, in Japan, it is necessary to examine how regulation should be applied to AI use not only in the private sector but also in the public sector, including the government. Furthermore, as the discussion on use of AI in defence and national security advances globally, discussions should be organised domestically, while taking account of the economic security policy perspectives. As an observer state, Japan should promptly confirm the existence of relevant domestic laws and guidelines, clarify desirable directions, and actively participate in drafting and negotiating the Convention.

Expeditious development of methodologies and systems for conducting risk and impact assessments is imperative

The present draft content of the Convention stipulates that measures should be taken for risk impact assessments to be conducted with respect to AI systems within the scope of the Convention. Therefore, a risk and impact assessment methodology and system that are feasible for relevant organisations should be developed in Japan as soon as possible, while considering trends in international standard risk and impact assessment frameworks. Currently, the Japanese government is in the process of formulating “Guidelines for AI Businesses”. While the contents of the Guidelines and this Convention need not be identical, the aim should be to ensure that there are no contradictions and that information on the Japanese Guidelines should be advocated.

Initiate immediately the discussions concerning the nature of the oversight mechanisms

The Convention is envisioned to mandate the establishment of an independent and impartial mechanism with the necessary authority and resources to monitor and oversee compliance with obligations. Although designating an existing institution as a structure or to using multiple institutions as an oversight mechanism is allowed under the Convention, concrete discussions should be initiated as soon as possible on how to create an efficient structure within the government, with an eye to the future, while referring to the responses of other countries.

The oversight mechanism will protect personal rights and play a role in asserting the legitimacy of Japan's efforts in other countries. Therefore, it is crucial to promptly consider and designate existing administrative organizations involved in the oversight mechanism and the number of personnel that are experts in AI should be increased and educated. In addition, experts should be promptly appointed to be involved in the negotiation stage of this Convention, and to hasten the establishment of a system that will receive support from the industry, academia, public, and private sectors. The main members of the oversight mechanism should be individuals who are capable of performing their duties over a long period of time with continuity, so that they can be recognised domestically and internationally as negotiating parties or contact points.

Process for Formulation of this Policy Recommendation

The content of this policy recommendation is based on the discussion in which the authors were organizing the issues on the Council of Europe's AI Convention. For interpretation of the legal perspectives, we received comments from Masahiro Ohya (Professor, Faculty of Law, Keio University), George Shishido (Professor, Graduate School of Law and Politics, University of Tokyo), Masahiro Sogabe (Professor, Graduate School of Law, Kyoto University), and Tatsuhiko Yamamoto (Professor, Law School, Keio University), and for the economic security perspective, we received comments from Toshiya Watanabe (Professor, Institute for Future Initiatives, The University of Tokyo). We also received comments from members of the Technology Governance Unit of the Institute for Future Initiatives, the University of Tokyo and support from Keisuke Shiraishi (Master's course, Graduate School of Public Policy, The University of Tokyo) during the compilation process. However, the authors assume responsibility for any factual inaccuracies or errors in this policy recommendation. Furthermore, we organized a webinar titled “The Current Status of the

Negotiation Process of the Council of Europe AI Convention and Experts' Perspectives”¹ hosted by the Institute for Future Initiatives, the University of Tokyo on July 11, 2023 and “Do you know about the AI Convention? -Discussions by the Council of Europe and Japan's Response” hosted by GLOCOM Roppongi Conference on August 4, 2023² and held discussions with a wide range of participants. In addition, relevant discussions were held in the “International Discussion on AI Strategy” panel discussion at the Japan Internet Governance Forum 2023 held on September 8, 2023³, as well as in a closed meeting on September 13, 2023 with overseas experts who participated in the Internet Governance Forum in Kyoto. Future plans include a panel discussion at the 23rd Information Network Law Association JAPAN Conference⁴ in December 2023.

¹ “The Current Status of the Negotiation Process of the Council of Europe AI Convention and Experts' Perspectives” hosted by the Institute for Future Initiatives, the University of Tokyo on July 11, 2023, <https://ifi.u-tokyo.ac.jp/project-news/16287/> (in Japanese)

² “Do you know about the AI Convention? -Discussions by the Council of Europe and Japan's Response” hosted by GLOCOM Roppongi Conference on August 4, 2023, <https://roppongi-kaigi.org/event/2816/> (in Japanese)

³ Japan Internet Governance Forum 2023 “International Discussion on AI Strategy” organized by Japan IGF, <https://japanigf.jp/meetings/japan-internet-governance-forum-2023> (in Japanese)

⁴ The 23rd Information Network Law Association JAPAN Conference, “The Council of Europe’s Negotiations on the Draft AI Convention and its Implementations for Japan,” <https://inlaw.jp/cf/23/#program> (in Japanese)

1. Commitment to International Collaboration on AI Governance

Considering the widespread adoption of artificial intelligence (AI) technologies, particularly machine learning and generative AI, and their potential to revolutionise various aspects of daily life and work, international discussions on AI governance have gained prominence. Organisations such as the Organization for Economic Co-operation and Development (OECD) and Global Partnership on AI (GPAI) are actively engaged in these deliberations. Furthermore, the G7 Hiroshima AI Process was established during the 2023 G7 Hiroshima Summit to harmonise diverse regulatory frameworks related to AI and promote collaborative efforts in AI governance. In addition, the Ministerial Declaration of the G7 Hiroshima Summit of the Ministers of Digital and Technology acknowledges the need for agile and multi-stakeholder governance and legal frameworks designed to operationalise the principles of the rule of law, due process, democracy, and respect for human rights.

The Council of Europe's (CoE) "The Framework Convention on Artificial Intelligence, Human Rights, Democracy, and the Rule of Law" (hereafter referred to as the AI Convention) is the world's first AI convention, that discusses a framework for considering values such as the rule of law, human rights, and democracy throughout the AI life cycle, from design, development, and use through to the decommissioning of AI. The CoE envisions the Convention as a framework agreement that welcomes the participation of democratic nations, including observer countries and non-CoE member states. All the G7 nations participated in the Convention's drafting negotiations. In negotiations, the United States and other observer countries, as well as Japan, emphasised compromise and shared a common awareness of the need to create a convention that facilitates the participation of many countries.⁵ The CoE actively encouraged non-CoE member states to participate in the negotiations, and the application of Argentina, Costa Rica, Peru, and Uruguay as observers of the Committee on Artificial Intelligence (CAI), which discusses the AI Convention, was adopted at the CoE ministerial deputy meeting on October 4, 2023.⁶

Considering the likelihood of many countries signing the AI Convention, it is imperative that Japan contemplates its approach. Furthermore, as the sole Asian observer country, Japan holds a position where it can provide input on the draft during the formulation stage and influence rule-making. In view of these circumstances, this policy recommendation indicates that Japan should confirm in the negotiation process an AI Convention to contribute to international discussions. Simultaneously, it seeks to stimulate domestic debates on AI governance and provides recommendations for Japan's negotiation process and post-signing and ratification responses, considering the complexity of this international landscape and its implications for Japan's AI strategy.

2. History of the Council of Europe's negotiations on the drafting of the AI Convention

Council of Europe

The Council of Europe (CoE) is an international organisation that upholds values such as human rights, democracy, and the rule of law. Established in 1949, it comprises 46 member

⁵ On September 18, 2023, Secretary General of the Council of Europe, Peččinović Burić called on countries to join Convention negotiations at a side event for foreign ministers during the UN General Assembly High Level Week, <https://www.coe.int/en/web/secretary-general/-/unga78-side-event-%C2%A0ministerial-event-on-artificial-intelligence-for-accelerating-progress-on-the-sustainable-development-goals-%C2%A0addressing-society-s-greatest-challenges>

⁶ 1477th meeting of the Ministers' Deputies, 04 October 2023, decision CM/Del/Dec(2023)1477/10.1b https://search.coe.int/cm/pages/result_details.aspx?objectid=0900001680accb1a

countries, including 27 countries of the European Union (EU).⁷ Russia was also a member until the invasion of Ukraine in March 2022 and included countries such as the United Kingdom, which left the EU in 2020, and non-EU countries such as Switzerland and Turkey. In addition to these 46 member states, the Holy See, United States, Canada, Japan, and Mexico participated as CoE observers. Israel is also an observer at CAI.

CAI and AI Convention

The CoE initiated discussions on AI in 2016. The need for AI to respect human rights became evident, leading to the establishment of an ad hoc Committee on Artificial Intelligence (CAHAI) within the CoE's Committee of Ministers in 2019. The purpose of the CAHAI was to explore the feasibility of creating a legal framework for the design, development, and use of AI that aligns with values such as human rights, democracy, and the rule of law. After two years of discussion, CAHAI prepared the groundwork for the creation of a legal framework, and in April 2022, the Committee on Artificial Intelligence (CAI) was established to draft an AI convention. CAI recognised the necessity for a legally binding document, leading to the initiation of discussions for the world's first AI convention titled "The Framework Convention on Artificial Intelligence, Human Rights, Democracy and the Rule of Law".⁸

Currently, the CoE is conducting intensive drafting negotiations and a Revised Zero Draft was released in January 2023. Subsequent negotiations on individual sections began at the CAI Plenary in July 2023, leading to the publication of a Consolidated Working Draft reflecting the negotiation outcomes. Further negotiations on the content were planned between autumn 2023 and early 2024 with the aim of adoption by the Committee of Ministers, the highest decision-making body, in May 2024.⁹

Column: The Council of Europe and the European Union

Table 1 compares the CoE and the EU, two international organisations located in Europe. These entities differ in their founding objectives, number of member countries, and legal documents. Regarding AI policies, the EU proposed a regulatory framework for AI in 2021. In June 2023, an amendment to the AI Act¹⁰ was passed, adding regulations on generative AI and currently underway on the finalization. By contrast, as mentioned earlier, the CoE aims to adopt the AI Convention in the first half of 2024.

One significant distinction is the scope of application. The EU's AI Act is limited to EU member states, whereas the CoE's AI Convention aspires to be a framework convention that includes democratic nations, including observer countries and non-CoE member states. Additionally, the legal layers of multilateral conventions and regulations differ between the CoE's AI Convention and EU's AI Act. However, if the EU became a party to the AI Convention, the AI Act could assume the role of implementing it. Although the contents of both documents do not need to be entirely identical, it is reasonable to assume that European countries aim to ensure that the AI Convention's content does not conflict with the AI Act.

⁷ The overview on the AI Convention presented in this paper was prepared with reference to presentations "AI Convention Negotiations Summary" and "Event report" of the "Current Status of the Negotiations on the Drafting of the Council of Europe AI Convention and Experts' Perspectives" held on July 11, 2023 by the AI Governance Project, Center for Future Vision Research, University of Tokyo. <https://ifi.u-tokyo.ac.jp/project-news/16287/>, in Japanese.

⁸ Council of Europe, CAI: <https://www.coe.int/en/web/artificial-intelligence/cai>

⁹ Revised Zero Draft and Consolidated Working Draft are available from the website: <https://www.coe.int/en/web/artificial-intelligence/cai>

¹⁰ EU AI Act: first regulation on artificial intelligence, European parliament, <https://www.europarl.europa.eu/news/en/headlines/society/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence>

Furthermore, given that a significant proportion of CoE member countries are EU member states, the EU's stance is crucial in the context of the AI Convention. Especially, concerning matters falling within the exclusive competence of the European Union, the European Commission is authorised to negotiations in the CAI in representing individual member countries.

Table 1 The Council of Europe and the European Union

	Council of Europe: CoE	European Union: EU
Year established	1949	1993
Founding objectives	Human rights, democracy, rule of law	Economic and political integration
The number of countries	46	27
Japan’s position	Observer	None
Main institutions	Committee of Ministers, Parliamentary Assembly, Congress of Local and Regional Authorities, European Court of Human Rights	European Council, European Commission, European Parliament, European Court of Justice
Legal documents	Multilateral Conventions, Recommendations to Member States	Regulations, Directives
AI policy	AI Convention drafting begins Guidelines and recommendations have been published	European AI act passed AI ethical guidelines have been published

3. Comprehensive Structure and Key Points of the AI Convention

Overview of AI Convention

The AI Convention is currently undergoing negotiations based on the Consolidated Working Draft released in July 2023. Consequently, its contents may be subject to further revisions in the future. At present, the Convention is structured and contains content as outlined in Table 2.

Table 2 Structure and Contents of the AI Convention

Preamble	
Chapter I : General provisions (Article 1 – 4)	
	Purpose and object, Risk-based approach, Definitions of AI systems and scope
Chapter II: General obligations (Article 5 – 6)	
	Respect for human rights and fundamental freedoms, Integrity of democratic processes and respect for rule of law
Chapter III: Principles of design, development, use and decommissioning of AI systems (Article 7 – 12)	
	Exceptions to these principles (e.g., national security), Principles such as transparency and oversight, accountability, equality, privacy, and safety
Chapter IV: Remedies (Article 13 – 14)	

	Ensuring the availability of effective remedies for violations of human rights and fundamental freedoms, Procedural safeguards (e.g., right to know)
Chapter V: Assessment and mitigation of risks and adverse impacts (Article 15 – 16)	
	Risk and impact management framework, Training
Chapter VI: Implementation of the convention (Article 17 – 22)	
	Non-discrimination, Rights of persons with disabilities and of children, Public consultation, Digital literacy and skills
Chapter VII: Follow-up mechanism and cooperation (Article 23 – 25)	
	Conference of the parties, International co-operation, Effective oversight mechanisms
Chapter VIII: Final clauses (Article 26 – 34)	
	Effects of the Convention, Amendments, Dispute settlement, Signature and entry into force, Accession, Territorial application, Reservations, Denunciation, Notification

As a fundamental premise, this Convention is a framework agreement, and while it possesses a legally binding force, the choice of means to achieve the objectives stipulated in the Convention is, in principle, left to the discretion of the contracting parties¹¹. Furthermore, it is assumed that the human rights protected by existing conventions also apply to AI.

The definition of AI in this Convention is provided in Article 3 (hereafter, unless otherwise specified, the article numbers refer to the Consolidated Working Draft). It defines AI as “any algorithmic system or a combination of such systems that uses computational methods derived from statistics or other mathematical techniques and that generates text, sound, image, or other content or either assists or replaces human decision-making”. It is explicitly stated that this definition includes generative AI, which is not mentioned in the Revised Zero Draft. Additionally, to adapt to technological advancements in AI, the Convention allows the contracting parties' assembly to “decide to give interpretation to this definition in a manner consistent with relevant technological developments”.

Furthermore, as a basic principle, this convention establishes measures for risk and impact assessments and advocates a risk-based approach. In cases of human rights violations, remedies are to be provided, and appropriate measures to guarantee AI system-related usage records for this purpose. In addition, it permits requests for human review when deemed appropriate. In cases where the evaluation results indicate outcomes incompatible with values such as human rights, actions such as the suspension of suspension of usage or mitigation of effects may be required. Furthermore, the Convention specifies the establishment of independent oversight mechanisms to monitor these provisions.

Key Points in the Drafting and Negotiations of the AI Convention

The AI Convention is currently undergoing negotiations, presenting challenges in defining the requisite actions for each country to adhere to the Framework Convention. Furthermore, a critical aspect of interpretation involves discerning whether complying with the Convention mandates the enactment of new legislation or whether existing legal frameworks, regulations, and guidelines can effectively address these concerns.

Therefore, by interpreting the disparities in content between the Revised Zero Draft and the Consolidated Working Draft, this policy recommendation aims to summarise three aspects that required in-depth scrutiny: (1) the Convention's scope of application, including

¹¹ A similar convention structure can be found in the United Nations Framework Convention on Climate Change.

exceptions; (2) measures for risk and impact assessment; and (3) oversight mechanisms. Items pertaining to any of these concerns are denoted by an asterisk, signifying a lack of consensus, and warranting vigilant monitoring of future negotiation developments.

(1) Convention's scope of application, including exceptions

Application to the Public and Private Sectors

Article 4.1 of the Consolidated Working Draft outlines that “This Convention shall apply to design, development, use and decommissioning of artificial intelligence systems that have the potential to interfere with the respect for human rights and fundamental freedoms, the functioning of democracy and the observance of rule of law”. Therefore, AI systems designed and employed by the public sector, such as the police and judiciary, are inherently within the scope of the Convention. Furthermore, the Convention extends to private sector involvement in public procurement.

The pivotal issue is whether the AI systems used in the private sector should be assessed and monitored for potential human rights violations. While Article 4.2 of the Revised Zero Draft explicitly states that the Convention would encompass both the public and private sectors, the Consolidated Working Draft allows for various interpretations by each party based on their respective constitutional and legal frameworks. Consequently, depending on the interpretation, private-sector entities may also fall under the Convention’s scope.

Additionally, Article 4.2 of the Consolidated Working Draft provides exceptions to the scope of application, excluding research and development activities unless the systems are tested or otherwise used in ways that have the potential to interfere with human rights. Nevertheless, private-sector research and development activities could also be effectively affected by the fact that the public sector, which allocates research and development funds, is bound by this Convention.

Defence and National Security as Exceptional Provisions

In the Revised Zero Draft, Article 4 defines the scope and explicitly states that “the Convention shall not apply to the design, development, and application of artificial intelligence systems used for purposes related to national defense.” However, in the Consolidated Working Draft, defence, along with several other areas such as national security, is categorised as an exception to the principles at the chapeau of Chapter III. If a provision had been incorporated into Article 4, it could have been construed as an exception to the Convention’s entirety. Nevertheless, its placement at the chapeau of Chapter III suggests that AI systems related to defence are solely exempt from the provisions contained in that chapter. The negotiating parties’ intentions in this regard remain ambiguous.

Even if the exceptions at the chapeau of Chapter III could be considered exceptions encompassing the entirety of this Convention, it is debatable whether defence as well as national security can be positioned as exceptions to the Convention. This ambiguity arises from the concern that surveillance activities that infringe on human rights may be conducted under the pretext of national security.

The use of AI for both civilian and military applications, particularly lethal autonomous weapons systems (LAWS), has been the subject of distinct deliberations. It is imperative to closely monitor how the Convention positions AI in the realms of national security and defence in forthcoming negotiations.

Exceptional Requirements

In addition to encompassing national security and defence, the chapeau of Chapter III delineates exceptions for various matters, including the protection of public safety, health and

morals, important economic and financial interests of the State, the impartiality and independence of the judiciary, or the prevention, investigation, and prosecution of disorder or crime, as well as the protection of the rights and freedoms of others. These exceptions are permitted solely if they are “already permitted under the domestic law of, and international legal obligations”.

Interpreting this requirement is a crucial point of contention. For instance, if the objective of restricting individuals' actions as a countermeasure against factors such as pandemics or terrorism is considered to contribute to public safety, health and morals, crime prevention, or related objectives, then AI systems associated with human rights, democracy, and the rule of law can be excluded from the scope of the Convention if they are “already” permitted under the domestic law of international legal obligations. However, if the AI system is not “already” recognised under domestic law at the Convention's entry into force, it may be necessary to develop new domestic laws, and depending on the circumstances, exceptions related to defence, national security, and similar areas may not be recognised. Therefore, a comprehensive discussion concerning the interpretation of the term “already” is imperative.

(2) Measures for risk and impact assessment

Article 15 of the Consolidated Working Draft mandates the adoption of a risk-based approach to identify, assess, prevent, and mitigate the risks and impacts on human rights, democracy, and the rule of law arising from the design, development, use, and decommissioning of AI systems within the scope of the Convention.

While Article 15 does not prescribe a specific methodology, the Convention's negotiating body, the CAI, is concurrently engaged in developing a risk and impact assessment methodology. It is essential to note that this methodology is not legally binding but serves as a model for parties to conduct risk and impact assessments in accordance with the Convention.¹² Parties have the flexibility to adopt existing risk and impact assessment methodologies, regardless of whether they carry legal force, provided they meet the requirements outlined in Article 15.¹³

Nevertheless, because the proposed CAI methodology is still in the development phase, further progress is anticipated. Some parties may employ existing risk impact assessment frameworks, such as those provided by the U.S. National Institute of Standards and Technology (NIST). However, it is crucial to recognise that the methodology itself should adapt in an agile manner to keep pace with evolving technologies.

¹² The guidelines are named “Human Rights, Democracy, and the Rule of Law Risk and Impact Assessment,” referred to by the acronym “HUDERIA.” The HUDERIA assessment process typically follows these steps: The first step is to conduct a risk and impact assessment based on the context of the AI use in question. If no risk to human rights is identified, the process concludes without further steps. Conversely, if there is a determined risk, stakeholders are identified, additional risk assessments are conducted with their involvement, and measures to mitigate the impact on human rights are taken. These processes must be repeated until the AI system in question is no longer in use. The Alan Turing Institute of the United Kingdom played a significant role in developing HUDERIA, including publishing the foundational report. The report submitted by the Institute to CAHAI as a draft of HUDERIA is available on its website. Human Rights, Democracy, and the Rule of Law Assurance Framework for AI Systems: A proposal prepared for the Council of Europe’s Ad hoc Committee on Artificial Intelligence

<https://rm.coe.int/huderaf-coe-final-1-2752-6741-5300-v-1/1680a3f688>

¹³ Article 15 also states “Each Party shall take measures for~” and uses the word “measures,” without regard to whether they are legally binding or not.

(3) Oversight mechanisms

Article 25 of the Consolidated Working Draft requires that a mechanism be established to oversee and supervise compliance with the obligations of the Convention. Each party shall ensure that such mechanisms exercise their duties independently and impartially, and that they have the necessary powers, expertise, and resources.

The point of contention here is that the use of the words “oversight mechanisms” instead of “independent supervisory authorities”, of the European General Data Protection Regulation (GDPR). This word allows for flexibility other than the establishment of an authority.¹⁴ While the GDPR's “independent supervisory authorities” provide detailed provisions on everything from the status of their members to the authority of the authorities, the Convention allows for the designation of existing authorities as supervisory mechanisms or for multiple authorities to be supervisory mechanisms under the Convention to the extent that they meet the Convention's requirements.

In the case of the GDPR, which requires the establishment of a domestic administrative organisation under a Convention originating in Europe, Japan, as a non-party country, has voluntarily established an administrative organisation that is recognised by the EU as being above a certain level to obtain a so-called “adequacy decision” to facilitate data transfers with the EU region. However, because the provisions of Article 25 of this Convention are expected to be directly legally binding to Japan as a contracting party, the establishment of an oversight mechanism is expected to be imposed on contracting parties as an obligation under international law. In other words, the establishment of an oversight mechanism is a condition for concluding this Convention.

4. Policy recommendations to the Japanese Government

Drawing upon the summarised discussions in the preceding sections, this section offers recommendations on the negotiation process if the Japanese government were to conclude this Convention, and on the measures to be taken if it were to ratify the Convention.

AI-related regulations that would be included in the scope of this Convention should be identified, and proactive participation in the negotiation process is essential

Japan's approach to dealing with cross-cutting disciplines in the development and use of AI systems has predominantly relied on guidelines rather than legislative measures, resulting in a scarcity of existing domestic laws. In addition, the applicable relationship with existing laws is often unclear. As previously highlighted in the exception provision at the chapeau of Chapter III, the question arises as to whether, at the time of the Convention's entry into force, the requirement for domestic laws (or international legal obligations) rather than guidelines, poses a substantial challenge for Japan in enacting new legislation aimed at encouraging the use of AI systems. This challenge persists even in cases where such systems are intended to enhance public safety but entail a notable risk of privacy infringement. This necessitates careful consideration to ensure that any new legislation aligns with the provisions of the Convention and avoids potential conflict.

Specifically, numerous terms like “domestic law” and “domestic legal framework” are

¹⁴ On the other hand, the Consolidated Working Draft states in Article 12 that “for a controlled regulatory environment for testing artificial intelligence systems under the supervision of its competent authorities.” It is assumed that this remains because CoE forgot to unify the word “oversight mechanisms” in Article 25, which suggests that there was some discussion during the drafting conception that called for authorities rather than mechanisms.

recurrently employed throughout the Consolidated Working Draft. During the negotiation process of the Convention, it is imperative to ascertain whether these terms carry distinct meanings and whether it is permissible to encompass guidelines within the domestic legal system (if not, negotiations could be made to change the wording to such a domestic legal framework to include guidelines). Additionally, deliberations on the interpretation of the term “already” hold significance within the Convention's negotiation process. It is advisable for active engagement in discussions concerning the interpretation of “already” and, when deemed necessary, advocating for its removal.

Moreover, because the use of AI in defence and national security is advancing globally, Japan also needs to conduct comprehensive domestic discussions on AI-related defence and national security. A summary of issues released in 2023 by the AI Strategy Council operating under Japan's Integrated Innovation Strategy Promotion Council note that “there are arguments that the use of AI may be important in security-related issues, but it should be handled flexibly, for example by leaving it for discussion by specialized divisions, depending on the need for information management”,¹⁵ and the practical situation is that security-related matters have not been adequately resolved on the domestic front. Considering the growing importance of preventing economic activities that harm the security of a nation and its citizens, the Economic Security Promotion Act was passed in 2022. The objective was to formulate a basic policy to advance economic strategies aimed at safeguarding security and constructing a comprehensive four-pillar system¹⁶. In parallel, the potential implementation of a security clearance system as part of the economic security measures to prevent the unauthorized export of sensitive technology is under evaluation.¹⁷ Consequently, an assessment of the interplay between the specific provisions of these laws and systems and their relationship with the Convention is required.

By leveraging its status as an observer country with the ability to engage in drafting and negotiation phases, Japan should ascertain the existence of pertinent domestic laws and guidelines, expeditiously determine its preferred course of action, and actively engage in the negotiation process.

Expeditious development of methodologies and systems for conducting risk and impact assessments is imperative

As highlighted in the preceding section, the Convention's scope of application may potentially be construed as limited to the public sector, encompassing the private sector solely in relation to its procurement activities. Nevertheless, in Japan, given that the courts are bound by the Convention, there is a possibility that the judiciary, as part of the public sector, might refer to the Convention as a guiding framework for interpreting domestic laws and regulations when disputes involving private sector entities escalate into litigation. Furthermore, if it is definitively established that the private sector falls within the scope of this Convention, its sphere of influence will expand significantly.

Consequently, it is imperative to establish a feasible methodology and system for conducting risk and impact assessments that are applicable to both public and private sectors. Notably, Article 13 of the Consolidated Working Draft mandates remedies in cases of human rights violations and necessitates that the relevant appropriate measures to ensure that the

¹⁵ AI Strategy Council operating under Japan's Integrated Innovation Strategy Promotion Council, Tentative summary of AI issues, https://www8.cao.go.jp/cstp/ai/ai_senryaku/2kai/ronten_youshi_eiyaku.pdf

¹⁶ Ministry of Economy, Trade and Industry, Outline of the Economic Security Promotion Act, <https://www.japaneselawtranslation.go.jp/outline/75/905R403.pdf>

¹⁷ Cabinet Secretariat, Expert Committee on Security Clearance System, https://www.cas.go.jp/jp/seisaku/keizai_anzen_hosyo_sc/index.html (in Japanese).

relevant usage of the AI system is recorded since the absence of such records could hinder the remedy process. From this perspective, it is essential to contemplate a risk and impact assessment methodology that considers feasibility.

As of 2023, Japan is actively engaged in the development of the “Guidelines for AI Businesses”,¹⁸ aimed at consolidating Social Principles of Human-Centric AI by the Cabinet Office,¹⁹ the AI Development and Utilization Guidelines by the Ministry of Internal Affairs and Communications (MIC)²⁰ and the Governance Guidelines for the Implementation of AI Principles²¹ by the Ministry of Economy, Trade and Industry (METI). The initial framework of this new guideline aligns with the principles outlined in the AI Convention, emphasising that every entity bears the responsibility of designing and employing AI systems in a manner that upholds “the rule of law, human rights, and democracy”, and underscores the adoption of a risk-based approach. To avoid redundancies in the risk and impact assessment process of domestic guidelines and the Convention subsequent to the Convention's ratification, it is crucial to closely monitor the assessment methodology being developed by the CAI. Moreover, it is advisable to disseminate information concerning the Japanese Guidelines to the negotiating table as deemed necessary.

Initiate immediately the discussions concerning the nature of the oversight mechanisms

Under the Convention, it is considered that parties are obligated to concurrently determine and formulate the framework of domestic AI-related regulations, devise methodologies and systems for conducting risk and impact assessments, and establish oversight mechanisms to ensure compliance with these regulations no later than the Convention's entry into force.

Regarding the actual organisational structure, if a new oversight body is to be established, the establishment of an independent administrative commission²² with the requisite independence and expertise mandated by the Convention is one option to consider. However, it is worth noting that the Convention does not preclude the possibility of designating an existing relevant administrative body as an oversight mechanism. Therefore, it may be possible to conceive a system wherein the supervisory authority is vested in an existing administrative agency with the necessary expertise, supplemented by a court responsible for adjudicating appeals against its decisions in legal proceedings. This hybrid approach can be recognised as a valid oversight mechanism under the Convention, ensuring both expertise and independence in a broader sense, encompassing the judicial branch within the Japanese government.

In addition, this oversight mechanism is anticipated to protect personal rights and to serve as a representative of national interests, asserting the legitimacy of Japan's endeavours to its counterparts from other parties. Therefore, it is crucial to promptly consider and designate existing administrative organizations involved in the oversight mechanism and the

¹⁸ The Cabinet Office, Skelton of “Guidelines for AI Businesses,” https://www8.cao.go.jp/cstp/ai/ai_senryaku/5kai/gaidorain.pdf (in Japanese).

¹⁹ The Cabinet Office, Social Principles of Human-Centric AI by the Cabinet Office, <https://www.cas.go.jp/jp/seisaku/jinkouchinou/pdf/humancentricai.pdf>

²⁰ Ministry of Internal Affairs and Communications, Draft AI Development Guidelines for International Discussion, AI Network Society Promotion Council, https://www.soumu.go.jp/main_content/000507517.pdf; Ministry of Internal Affairs and Communications “AI Utilization Guidelines,” https://www.soumu.go.jp/main_content/000658284.pdf

²¹ Ministry of Economy, Trade and Industry, Governance Guidelines for the Implementation of AI Principles, https://www.meti.go.jp/shingikai/mono_info_service/ai_shakai_jisso/pdf/20220128_2.pdf

²² Refers to a commission that is an external department of a ministry or agency as defined in Article 3, Paragraph 2 of the National Government Organization Act (Including Article 49, paragraph 1 of the Act for Establishment of the Cabinet Office, which provides to the same effect).

number of personnel that are experts in AI should be increased and educated. In addition, experts should be promptly appointed to be involved in the negotiation stage of this Convention, and to hasten the establishment of a system that will receive support from the industry, academia, public and private sectors. The main members of the oversight mechanism should be individuals who can perform their duties over a long period of time with continuity, so that they can be recognised domestically and internationally as negotiating parties or contact points.

Concrete discussions should be initiated promptly considering the organisational requirements stipulated in the Convention and the nature of the oversight mechanisms essential for AI governance in Japan. It is also advisable to refer to the approaches adopted by other nations in this regard.

5. International significance of this Convention and the prospects for Japan's AI governance strategy

This policy recommendation offers an interpretation of the Consolidated Working Draft of the AI Convention and outlines the measures that Japan should take. AI services and systems are continuously being developed, deployed, and used in various countries and regions. In an era marked by a shifting balance of power among regions, including Europe, the U.S., the Global South, and East Asia, coupled with emerging geopolitical challenges, the pivotal question revolves around the values that govern the development and use of AI. In this context, the CoE aspired to establish the Convention as a shared framework for countries within a democratic sphere. We hope that this recommendation will contribute to the future drafting and negotiation of the AI Convention by the Council of Europe, as well as to the formulation of Japan's AI strategy.