Towards an AI governance framework (ver2.0)

Dec 2024.

Digital Policy Forum Japan

Basic concept

Technological developments around generative AI are progressing at a remarkable rate (Ref. 1)[1] and the implementation of generative AI in socio-economic systems is accelerating (Ref. 2).

In this context, the development of rules to control AI is rapidly steering from ideological discussions to concrete ones. Some countries and regions have passed AI-related legislation, such as the European "AI Act"[2] and China's "Interim Measures for the Management of Generated AI Services"[3], while in the US, specific governance rules for AI are being discussed at various levels[4], and in Japan, discussions including the introduction of legislation[5] are about to start in earnest.

In this document, keeping in mind these developments around generative AI, the basic perspectives of the study are,

- Minimizing the risks of AI,
- Develop an environment that maximises the convenience of AI, and,

---

1 Document numbers are excerpt from the annex 'Trends in AI' (same below).

2 In May 2023, the European Council adopted the 'AI Act', which entered into force in stages from May 2024, with full application scheduled for summer 2026.
https://artificialintelligenceact.eu/

3 In August 2023, China implemented the Regulations for the Management of Generated Artificial Intelligence Services. The Regulation (Article 4) prohibits the creation of content prohibited by law or administrative regulations and only allows generated material that 'adheres to the core values of socialism'.
(Source: Masashi Harada, 'China's "Interim Measures for the Management of Generated Artificial Intelligence Services" and its Commentary', Corporate Legal Navigator (21 July 2023).
https://www.corporate-legal.jp/matomes/5362

4 In October 2023, the US Government published a Presidential Decree on AI governance. As measures to be undertaken by government agencies, the order includes the establishment of standards for vulnerability research (Red Teaming), clear guidance on the prohibition of algorithmic discrimination, and support for the appropriate use of AI in healthcare, education and other sectors. However, President-elect Trump has already announced his intention to repeal this Presidential Decree. See footnote 10 for information on the public-private partnership initiatives that preceded the Presidential Order.
https://www.whitehouse.gov/briefing-room/statements-releases/2023/10/30/fact-sheet-president-biden-issues-executive-order-on-safe- secure-and-trustworthy-artificial-intelligence/

5 See, for example, AI Strategy Council, 'On the "Approach to AI Institutions"' (May 2024).
https://www8.cao.go.jp/cstp/ai/ai_senryaku/9kai/shiryo2-1.pdf

- Creating a generative AI market that makes such an environment as autonomous as possible.

This section provides a direction for consideration of the establishment of an 'AI governance framework', or in other words, a mechanism to continuously maintain the 'controllability of AI technology', in order to realise the three objectives in a balanced manner.

AI governance needs to be discussed by constantly weighing the balance between the benefits and risks that AI brings. AI contributes to improving productivity and creativity in all areas of society, and it brings a wide variety of benefits, such as being able to enjoy highly convenient services while technically ensuring data sovereignty (data sovereignty) for the use of personal data through personalization (decentralization of intelligence). On the other hand, there are concerns about the increasing severity of human rights violations, the risk that humans will be unable to control AI technology, and the risk that AI will replace humans. Incidentally, in order to promote innovation, it is not appropriate to hastily introduce regulations, and it is thought that it is preferable to resolve these risks as much as possible through technical means.

The discussion in this document will be conducted primarily with generative AI that is currently available to the general public in mind, and will not cover Artificial General Intelligence (AGI), with some exceptions.

In July 2024, the Digital Policy Forum (DPFJ) published this document as ver 1.0[6] , which summarised the main issues surrounding AI governance, and has since conducted interviews with experts and interviews with business stakeholders, as well as organising additional issues. Current ver 2.0 clarifies the basic direction of the main issues as far as possible. However, these basic directions are only a draft, and necessary amendments will be made according to the progress of future discussions on AI governance, and in light of the rapid evolution of AI technologies.

I Risk minimization.

1. Risk management

a) Difficulties in managing AI-related risks at each stage

There are several methods of AI management that divide risks (including negative impacts on human life and basic human rights) into several levels. For example, the EU AI Act (Ref. 3) classifies risks into four levels[7]. This is a system of managing the risks

[6] https://www.digitalpolicyforum.jp/wp-content/uploads/2024/06/240701_AI01.pdf

[7] The European AI Act classifies AI risks into four categories: unacceptable risk (development prohibited as posing a direct threat to human life or fundamental human rights), high risk (obligation

associated with AI models according to their severity and linking them to the degree of regulation. However, in the case of this approach, in addition to specific risk management methods, such as how to define the scope of risks to be controlled and what criteria to rank the risks, it cannot be said that the entities responsible for making risk decisions and the methods for clearly indicating to third parties the accuracy of the decisions made by these entities (accountability) have been established.[8]

It should be noted that the sources of risk associated with AI are diverse, and it is difficult to ascertain the full extent of the risks. For example, according to the MIT survey (Ref. 4)[9] , there are more than 700 risks associated with AI, and it is very difficult to implement a risk management system with all of them in mind. It is also necessary to take into account the fact that risks change dynamically and qualitatively over time, as indicated in the survey, with 'post-deployment risks' accounting for 65% of all risks.

Of course, AI risk management itself is extremely important and, in Japan, industry-academia-government collaboration should actively promote the creation and analysis of repositories on AI risks.

b) Risk management by entity

While taking the above into account, it is advisable to consider risk management[10] of AI separately for three actors: developers of AI, service providers implementing AI and end-users.

The risk management by AI developers should be limited to a "Do Not List" approach, which lists a limited number of issues that need to be taken into account during development. In the future, if specific problems arise in AI development, they should be

---

to conduct prior conformity assessment, register in database, etc.), limited risk (obligation to ensure transparency to inform users that they are interacting with AI), and minimal risk (no regulation).

[8] In the future, if a system for scoring AI risks by recording and analysing AI system logs (operating history) is established and this information is made public, users may be able to select AI according to the risks they can tolerate. In other words, it is necessary to actively participate in discussions on AI standardisation (including risk assessment methods) in international organisations, bearing in mind that a mechanism can be established to enable each user to select a (personal) AI suitable for his/her own use by comparing and weighing the benefits (benefits) and risks (costs) of the AI in question. It is also necessary to actively participate in discussions on AI standardisation (including risk assessment methods) in international organisations.

[9] P. Slattery et al. "Global AI adoption is outpacing risk understanding, warns MIT CSAIL" (MIT CSAIL News, August 14, 2024).

[10] When considering the risks of AI, the risks that AI may entail at the development stage and the risks that AI may have at the service provision stage (risks that may be manifested by the way services implemented with AI are provided and used, e.g. the generation and dissemination of false information). For example, the generation and dissemination of misinformation and false information). In particular, with regard to the latter risk, it is necessary to carefully discuss whether the risk is brought about for the first time by AI, or whether it is a risk that has existed for a long time but has become apparent or amplified due to AI.

dealt with at the time of occurrence, while regular monitoring may be conducted.

Specifically, the principle could be, for example, to "ensure that activities in the lifecycle of AI systems are fully compatible with 'human rights, democracy and the rule of law' ", with reference to the Council of Europe's 'The Framework Convention on Artificial Intelligence' (September 2024) (Ref. 5)[11] , for example. The principles include.

Risk management by service providers implementing AI should also be as limited as possible. For example, as stipulated by Article 6 of the Telecommunications Business Act (Japan)[12] , this may be limited to such disciplines as prohibiting unfair and discriminatory treatment in the provision of services.

The prohibition of unfair discriminatory treatment in relation to AI is because, while AI makes it possible to provide detailed services using personal information, such as personalised medical care, it is required from the perspective of human rights protection to ensure that such services are not personalised according to individual characteristics and do not constitute discrimination lacking a rational basis (see item (5) for specific measures).

When providing users with services incorporating AI, the boundary of responsibility between the developer of the AI and the service provideris also required to be clarified in advance at the stage prior to the provision of the service, from the perspective of user protection.

Furthermore, for risk management at end-users (including SMEs, etc.), literacy education is required to ensure a correct understanding of the risks of AI[13] (see item (5)).

c) Risk management methods

Risk management should be based on the results of the preparation and analysis

---

11 In September 2024, the Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law was signed by 10 countries and regions, including the US and the EU (Japan has not signed). Rule of Law), signed by 10 countries/regions including the US and the EU (Japan has not signed).
https://www.coe.int/en/web/artificial-intelligence/the-framework-convention-on-artificial-intelligence

12 Article 6 of the Telecommunications Business Act states that "Telecommunications carriers shall not provide unfair and discriminatory treatment in the provision of telecommunications services." It stipulates that.

13 For example, in September 2024, an amendment to the Special Act against Punishment of Sexual Violence Crimes, etc. was passed in South Korea to punish the possession and viewing of sexual deep-fake images and videos. Also, at the state level in the US, 19 states have introduced labelling regulations for AI-generated content in the context of elections (as of July 2024*), and there have been developments in several countries regarding laws and regulations from the perspective of users (voters).
*(Source) Funk, Vesteinsson, Baker, Brody, Grothe, Agarwal, Barak, Loldj, Masinsin, Sutterlin eds. Freedom on the Net 2024, Freedom House (October 2024).
https://freedomhouse.org/report/freedom-net/2024/struggle-trust-online

of AI risk repositories, as already mentioned, and in doing so, consideration should be given to whether risk self-assessment or third-party assessment (e.g., audit or certification systems) should be applied.

Specifically, for developers of AI, this could be based on self-assessment by the developers themselves, combined with a third-party audit system for those that are closely related to socially essential critical services.

As it is difficult for service providers to extract and evaluate only the functions of AI, they should be considered and operated within the existing framework for user protection in each business law, and it is not appropriate to add additional regulations on the occasion of the use of AI.

2. Regulation and effectiveness

Regulatory approaches to AI include hard law (legislation) and soft law (self-regulation by the private sector), as well as co-regulation through public-private partnerships that combine these approaches[14] (Ref. 6). For example, in China (Ref. 7) and the EU, hard law is the basic[15] , while in the US (federal government), policy development has been based on private sector self-regulation (Refs. 8-10). However, even in cases where hard law is orientation, there is a certain range in terms of the nature of discipline, such as a loose basic legal approach and a highly disciplined approach that imposes specific conduct regulations.

Amid rapid technological innovation, some AI-related discussions in the past tended to be far removed from the realities of the market and tended to be more abstract and speculative than necessary. The basic principle should be to ensure the necessary discipline, promote the digital industry (appropriate balance between regulation and promotion), and realize international harmonization of discipline in a three-pronged manner, based on calm

---

[14] Joint regulation is a method whereby the State sets out the basic policy of the rules, operators who agree with the purpose of the rules operate the rules based on the basic policy and report the results to the State, which then evaluates the results and amends the basic policy as necessary. In Europe, this system has been adopted as a countermeasure against disinformation by platform operators. While co-regulation is excellent in terms of flexible application of rules led by the private sector, on the other hand, sufficient transparency is required to ensure that administrative rules are not applied in a discretionary manner without a legal basis.

As an example of public-private partnership, although not an example of co-regulation in the AI sector, a non-binding agreement was reached between the Office of the President and seven AI-related companies (Amazon, Anthropic, Google, Inflection, Meta, Microsoft & OpenAI) in July 2023, prior to the publication of the Presidential Decree (see footnote 3). In September of the same year, eight companies (Adobe, Cohere, IBM, Nvidia, Palantir, Salesforce, Scale AI and Stability) joined this agreement in addition to the above seven companies.

[15] It should be noted that in the EU, direct regulation through hard law prohibits "unacceptable risks" through legislation, while the overall system is designed with co-regulation of "high-risk" AI, etc. through codes of conduct, etc. in mind.

discussion and the voluntary efforts of the parties concerned.

a) Enactment of the AI Basic Act

When considering a legal system for AI in Japan, it is appropriate to enact the AI Basic Act as a hard law, rather than legislating in detail the contents of the various guidelines that have been studied by the Government so far, referring to, for example, the Cyber Security Basic Act[16] , It may stipulate the basic principles of policies on AI, the responsibilities to be fulfilled by the State and other actors, the formulation of AI strategies, the powers of the AI Strategy Headquarters (and the Secretariat of the Headquarters) in the Government, and cooperation with relevant organisations.

Cross-industry initiatives such as risk management by developers should be carried out mainly by the Headquarters Secretariat set up in the Cabinet Secretariat, while service providers should be carried out by the competent authorities for each business category, and in particular, when it is deemed necessary to ensure matters (unified standards) in a cross-industry manner from the perspective of user protection according to the characteristics of AI. In cases where it is recognised that unified standards are required, the Headquarters Secretariat should take the lead in promoting unified measures in cooperation with the relevant ministries and agencies.

b) AI Basic Act and the Role of the State

It is not necessarily appropriate to assume in the institutional framework that AI developers should be widely subject to legal control (registration or notification system). This is because sufficient rationale and public consensus are required for the legal regulation of AI development as a part of the development of science and technology in which anyone can participate.

From this perspective, for example, the self-audits and external audits by third parties expected of developers by the government should be voluntary measures by developers, and consideration could be given to establishing a mechanism for cooperation with the government (AI Strategic Headquarters) as necessary (for example, the government, etc. could formulate audit guidelines and revise such guidelines based on the actual status of the audit, etc.). The establishment of a mechanism (e.g., the government, etc. formulates audit guidelines and revises such guidelines based on the actual status of audits) may be considered. There have been

[16] The Cyber Security Basic Act stipulates the basic principles and responsibilities of each entity (national government, local governments, critical infrastructure providers, etc.) with regard to cyber security measures, the formulation of cyber security strategies, basic measures and the establishment of the Cyber Security Strategy Headquarters.
https://laws.e-gov.go.jp/law/426AC1000000104#Mp-Ch_1

some discussions on the introduction of regulations for developers of particularly large-scale AI, but as mentioned above, the introduction of voluntary external audits by a third party is only an option for large-scale AI, and should be clearly distinguished from competition policy-related discussions (see item (6)) on how the size of the market affects the relevant market. see item (6)) could be clearly distinguished from the competition policy related discussion on how size affects the relevant market.

3. Vulnerability measures against external risks

As AI becomes a social infrastructure, functional assurance (mission assurance) to ensure the resilience of AI[17] is extremely important. For this reason, all parties concerned need to work together, especially on measures to address external risks such as AI vulnerabilities.

a) Countermeasures against cyber-attacks related to AI

From the perspective of managing external risks of AI models, it is appropriate to incorporate AI vulnerability investigations (red teaming) into the audit (self-audit or external audit by a third party) items, and guidelines for implementing this should be developed through public-private partnership. In this regard, in Japan, the AI Safety Institute (AISI) published the Guide to Red Teaming Methodology for AI Safety (Version 1.00)[18] in September 2024. When considering this, it is extremely important to limit and clarify the scope (purpose) of vulnerability investigations, etc., in view of the wide range of characteristics of AI, from the perspective of ensuring their effectiveness.

There is also a risk that the AI may not perform its intended functions or malfunction due to data contamination attacks[19] and others during the AI learning process (cyber-attacks against AI). There is also an emerging risk of AI being used to discover vulnerabilities, create malware, generate fake accounts and distribute false information

---

[17] Functional assurance refers to "the process of defending and ensuring the ongoing functional maintenance and capability resilience of capabilities and assets required for (DoD's) Mission-Essential Functions (MEFs) --- personnel, equipment, facilities, networks, information and information systems, infrastructure and supply chains --- under any environment or condition" (source: US Department of Defense "Mission Assurance Strategy" (April 2012)). It refers to "the process for defending and ensuring the ongoing functional maintenance and capability resilience of the capabilities and assets required for the personnel, equipment, facilities, networks, information and information systems, infrastructure and supply chain" (Source: US Department of Defense "Mission Assurance Strategy" (April 2012)).

[18] https://aisi.go.jp/assets/pdf/ai_safety_RT_v1.00_ja.pdf

[19] In a data poisoning attack, an attempt is made to modify the model to function maliciously by inserting tainted data into the training data that produces incorrect outputs. In a data evasion attack, noise or other elements that cannot be perceived by humans are mixed into the training data to mislead the AI's decision-making results.

(cyber-attacks against AI). Specific measures to deal with such 'cyber-attacks against AI' and 'cyber-attacks by AI' also need to be urgently considered (Ref. 11).

In considering the above, it is necessary to simultaneously consider both the need to ensure openness and the possibility of AI being misused, from the perspective of avoiding the appearance of vulnerabilities and the malicious imitation or misuse of AI systems by third parties by making the learning data and AI systems open.

b) Ensuring the soundness of the data space

In the process where AI repeatedly learns from learning data, it often adopts a process of abstracting infrequently occurring data (with the aim of improving hit rate for queries). In this case, words with high occurrence probability in the previous generation model are valued in the next generation, while words with low occurrence probability are undervalued, resulting in a loss (degradation) of model diversity. This phenomenon, known as 'model collapse,' has been pointed out as a potential issue (Ref. 12)[20]. Leaving this situation unchecked will lead to the dissemination of inaccurate and unsound data and the ongoing contamination of the data space.

For this reason, it is necessary to consider the establishment of a private-sector-led certification system, for example, to limit AI learning data to those created by humans, or to clearly indicate to the outside world that the AI is a trained AI. In addition, from the perspective of increasing the amount of human-created data, it would be effective to make documents whose copyrights have expired, documents created by public institutions, etc. widely available as open data for use as learning data.

4. Handling of the generated products

AI takes in learning data, forms a model and utilises it to output data as a product. Therefore, from the perspective of ensuring data integrity, (3) above, 'ensuring the soundness of the data space' is from the perspective of ensuring the integrity of the input values (learning data), but at the same time it is also necessary to work to ensure the integrity of the output values (products). Therefore, in a situation where a vast amount of dis/misinformation is already circulating using generated AI, it is necessary to effectively and concretely promote countermeasures against disinformation while assuming a co-regulatory approach.

---

[20] I. Shumailov et al. "The Curse of Recursion: Training on Generated Data Makes Models Forget" arXiv (May 2023).
https://arxiv.org/abs/2305.17493

In such cases, the introduction of digital watermarking, which enables the identification of AI products, is considered effective. The effectiveness of originator profile (OP) technology, which enables users to confirm the creator and originator of information (content) on the Internet, also needs to be discussed in relation to the international standardisation of technical standards and the way the entity that grants OPs should be.

II Increased convenience

5. Active use of AI

a) Promoting the use of AI to solve issues

Various initiatives have already begun on the use of AI, but given that data utilisation efforts are lagging behind in the education and healthcare sectors[21] , especially in the context of a seriously declining birthrate and ageing population, it is necessary to actively promote the use of AI in these fields.

In particular, a system that links and analyzes relevant data under individual consent, starting with students in education and patients in healthcare, is expected to contribute to the individualisation of education and healthcare.

On the other hand, certain safeguard measures should also be considered to ensure that such data linkage does not lead to excessive profiling. In addition, AI analysis will enable automatic linkage of cases where data linkage has not progressed, for example, medical record data, due to differences in data formats between regions and organisations.

In addition to the fields of education and medicine, AI needs to be actively utilised in a wide range of other fields, such as environmental measures, which are global issues, disaster prevention and mitigation to protect human life and property, and culture to realise a prosperous life. In doing so, it is necessary to deepen consideration of the matters that need to be taken into account and the technologies that need to be developed in order to actively utilise AI in these fields.

At the same time, it is necessary to take the necessary measures to protect privacy, including the handling of personal data as learning data and avoiding the possibility of personal data being included in the output of when such data is imported. In addition, clarification is required on the treatment of learning data and generated materials under copyright law (Refs. 13-14).

---

[21] In the medical sector, for example, it is expected to prescribe therapeutic drugs based on personal data, predict disease risks and improve accuracy, realise rapid and efficient drug discovery, and automate medical administrative tasks.

In addition, as already mentioned, literacy education is important to ensure that general users correctly understand the risks of AI. For example, it is important to conduct public awareness-raising activities on the risks of AI, as in the case of efforts to improve the internet use environment for young people through public-private partnerships.

b) Promoting the use of AI in administrative services

In the provision of administrative services by the national and local governments, it is necessary to efficiently deploy limited human resources and provide more personalized and detailed services through the use of AI and active data collaboration, as the birthrate continues to decline and the population ages. However, it is essential to gain the understanding of local residents with regard to the active use of AI in the provision of such administrative services, and therefore the necessary institutional framework (formulation of basic guidelines and implementation of risk assessment) should be developed and put into operation, referring to the example of Kobe City, Hyogo Prefecture[22], efforts should also be made to share best practices.

c) AI and the labor market

Some argue that the active use of AI will lead to the automation of society and the loss of employment opportunities (i.e., the loss of human jobs). However, the basic policy is to utilise AI as a tool for improving labour productivity and creating new market areas, rather than aiming to use it to replace the existing workforce, and the Government is expected to provide the necessary policy support in the direction of achieving this.

Digital technologies, including AI, are not originally intended to promote efficiency in existing markets. Rather, it is necessary to share a broad understanding that it creates new employment by breaking down barriers in existing business areas and creating new market areas.

III Fostering Sound Markets

6. Building a sound ecosystem

The evolution of AI should basically be driven by the ingenuity of the private sector. The

[22] Ordinance on the use of AI in the City of Kobe (enacted in March 2024, entered into force in September of the same year).
https://www1.g-reiki.net/city.kobe/reiki_honbun/k302RG00001955.html

State should actively support this and provide the necessary rulemaking and policy support from the perspective of ensuring the public interest.

In doing so, competition policies to establish a sound market environment are important to ensure an ecosystem of diverse actors, including developers and users of AI.[23]

Therefore, it is necessary to establish a system to monitor anti-competitive behavior in the AI-related market, such as 'barriers to entry' and 'abuse of a dominant position by large companies. In addition, although the current major leading AI is mainly provided by existing large platform operators, the possibility of market dominance being abused in the AI market or adjacent markets (e.g. platform businesses) in the future and competition safeguard measures against this need to be considered.[24]

In particular, there is concern that vertically integrated AI developers with multiple layers of business development, such as platform operators, may have higher market dominance than other developers and are more likely to exercise market dominance over neighbouring markets, and it is necessary to consider how this should be addressed as competition policy.

In addition, the nature of market demarcation when conducting verification on whether there is an abuse of market dominance should be examined, with a view to cross-border distribution of data, networking of AI and AI collaboration across language barriers.

The European AI Act includes provisions for the extraterritorial application of laws, but consideration must also be given to the possibility that an increase in such extraterritorial application could lead to excessive regulations, such as the superimposition of foreign regulations on domestic ones.

7. Industrial promotion and global cooperation

The current use of computers is progressing towards optimizing the allocation of resources to meet diverse needs through a combination of centralized cloud computing and decentralized edge computing. Similarly, a world in which AI is networked beyond national borders is envisaged, where AI functions are enhanced by combining centralised and distributed computing resources and networked AI interactions. Given such a world, it will be essential to ensure openness of AI and globalisation of rules (see item (8)).

a) Ensuring openness

[23] OECD "Artificial Intelligence, Data and Competition" OECD Artificial Intelligence Papers No. 18 (May 2024).
https://www.oecd.org/daf/competition/artificial-intelligence-data-and-competition.htm

[24] Fair Trade Commission, Competition over Generative AI (Discussion Paper) (October 2024).
https://www.jftc.go.jp/houdou/pressrelease/2024/oct/241002_generativeai_02.pdf

One of the main reasons for the explosion of the Internet is its openness. Similarly, there are two possible approaches to AI: closed proprietary AI and open AI, but from the perspective of promoting healthy market development and maintaining the quality of AI-related services, ensuring openness to create a sufficiently competitive environment is essential. Similar approaches can be found in Europe and the USA[25].

In this context, the Government should actively promote the use of open source, how to ensure interoperability between different AIs, promote standardisation to create such an environment, and support research and development on the basis of encouraging open-type AI development.

As Japan is already lagging behind in the global market with regard to AI-related technological development, the Government should consider taking proactive measures to promote open-type AI, such as State support for the development of solutions incorporating open-type AI. In particular, discussions should be held to strengthen initiatives to support AI-related ventures.

In this case, we should distinguish between formal and substantive openness and ensure substantive openness in policy. For example, if the learning data for AI or specific feedback in the process of RLHF (Reinforcement Learning from Human Feedback) is not disclosed, there is a concern that substantive openness of AI will not be ensured (or proven) even though the openness of the technical specifications is ensured. In this way, safeguards to ensure substantive openness are also required.

b) Promoting a comprehensive AI strategy as an industry

The use of generative AI in Japan has remained partial within companies, and the number of cases that have led to business transformation is still limited; an AI-implemented industry means building a new, data-driven business model. Therefore, in formulating an AI strategy in government, it is required to formulate and promote an overarching comprehensive AI strategy that includes the development of relevant highly advanced technologies, semiconductor manufacturing and distribution, development of language models, environmental development for data distribution[26] , mechanisms

---

[25] In Europe, a list of issues relating to generative AI and competition policy was presented in the invitation document "Competition in Virtual Worlds and Generative AI: Calls for Contribution", published in January 2024. The list of issues related to generative AI and competition policy is presented.
https://ec.europa.eu/commission/presscorner/detail/en/ip_24_85
In addition, in the United States, the Presidential Decree (see footnote 3) lists 'promoting a fair, open and competitive ecosystem' as one of the main thrusts from the perspective of encouraging innovation and competition.

[26] Data Society Advancement Council (DSA), Digital Policy Forum (DPFJ) and Digital Trust Council (JDTF) Recommendation 'Promoting a Data Governance Strategy' (October 2024).
https://prtimes.jp/main/html/rd/p/000000009.000131931.html

for handling rights such as intellectual property and copyright, and other aspects of economic security.

8. Fostering international consensus.

It is assumed that AI will not be developed and used exclusively within a country, but will be networked and widely used in cyberspace. The above issues should be reflected in the legal systems and other rules of each country and harmonised as necessary, while forming a loose international consensus on the above issues.

In this context, given that AI is a strategic field and has a significant impact on industrial competitiveness and problem solving in each country, a bird's-eye view approach by experts in various fields such as industry, technology and diplomacy is required, and an effective system should be established within government departments and through public-private partnerships. In addition, given that AI has great potential to contribute to the resolution of issues faced by the Global South, it is necessary to proceed in a manner that involves the full participation of the Global South.

Furthermore, a particularly urgent task in fostering such an international consensus is the formation of norms for the military use of AI, as proposed at the Conference on Responsible AI in the Military Domain (REALM Summit) held in The Hague in February 2023, "Responsible Military Use of Artificial Intelligence and Autonomy. Political Declaration on the Responsible Use of AI" (Ref. 15), available at[27] , should be expanded to include voluntary commitments on the use of AI. At the same time, the inclusion of an AI security audit (inspection) mechanism within the UN security framework is worth considering. Discussions on the nature of such AI and security need to be hastened in light of the fact that the military use of AI has already become a reality (Ref. 16)[28].

---

[27] This proposal (US DoS "Political Declaration on Responsible Use of Artificial Intelligence and Autonomy" (February 2023)) proposes that military AI be be used only in a manner consistent with the obligations of international law (in particular, international humanitarian law); publish principles for the design, development, deployment and use of military AI; implement measures to minimise unintended bias; develop auditable military AI; and provide rigorous testing and assurance of the safety, security and effectiveness of military AI throughout its lifecycle. The content of the agreement includes voluntary commitments by states to conduct rigorous testing and assurance over the entire lifecycle of military AI, and 51 countries, including Japan, have now endorsed the agreement.
https://www.state.gov/political-declaration-on-responsible-military-use-of-artificial-intelligence-and-autonomy/

[28] According to an investigative report by Israeli online media outlet +972 Magazine in April 2024, the Israeli military is using a generative AI, Lavender, to extract 37,000 people in the Gaza Strip to make a list of operatives and target them, among other actions.
(Source) Yual Abraham "'Lavender': the Ai machine directing Israel's bombing spree in Gaza" +972 Magazine (April 3, 2024
https://www.972mag.com/lavender-ai-israeli-army-gaza/

9. Dealing with ethical issues.

With the rapid progress of AI, the possibility of 'self-conscious' AI in the future needs to be taken into account. Therefore, as in the life sciences, ethical issues related to AI research should be considered and specific research ethics codes and research approval processes should be established. For example, ethical guidelines for issues such as 'giving AI self-consciousness' and 'to what extent should it have the ability to self-replicate and modify itself' need to be developed and implemented.

Future work program

As indicated in the introduction, the basic theme of this document is 'Controllability of AI technology', in other words, the aim is to create an environment in which humans make the final risk decision and take responsibility themselves for the impact of AI.

The DPFJ will continue to update this document by holding workshops with relevant stakeholders on the basis of this document (update to ver 3.0 by summer 2025). At the same time, the document will be used as an opportunity to deepen discussions on the establishment of a broad AI governance framework, for example by holding open forums. In doing so, we will actively promote cooperation with other forums and other organizations that are pursuing similar discussions, in order to foster consensus.

See also Yasunori Kawakami, 'Targeting 37,000 people in Gaza: existence of AI machine 'Lavender' revealed', Yahoo! News (9 April 2024) for details of the above investigation.
https://news.yahoo.co.jp/expert/articles/c72d4cbc32aa5577eac494dfd75b43652a20555f