Japan's Response to New Technologies

Draft Artificial Intelligence Research and Development Guidelines for International Discussions

Souichirou Kozuka[*, **]

I. INTRODUCTION

Like other major economies, Japan is focusing on Artificial Intelligence (AI) as a key driver for economic growth. It is also expected that AI systems will be useful in providing solutions to social problems experienced by aging societies having a decreasingly young workforce. As a result, AI has become something of a "buzz word" in Japanese policy-making. Various government agencies are producing policy instruments referring to the use of AI.

Because AI systems will have a larger potential when they are connected to each other through a network, the use of AI is often associated with an extended network, referred to as the Internet of Things ("IoT"). The Japanese government has promulgated a concept of "Society 5.0" and has declared in the Fifth Science and Technology Basic Plan that it stands as Japan's goal for the future. "Society 5.0" is a "human-centered society that balances economic advancement with the resolution of social problems by a system that highly

* Professor, Faculty of Law, Gakushūin University, Tōkyō.
** The research for this paper has been funded by the "subsidies for scholarly study of non-life insurance" from the General Insurance Institute of Japan.

integrates cyberspace and physical space."[1] According to the Japanese government, in Society 5.0 "a huge amount of information from sensors in physical space is accumulated in cyberspace. In cyberspace, this big data is analyzed by artificial intelligence ("AI"), and the analysis results are fed back to humans in physical space in various forms".[2]

Given that the network relies on both wired and wireless telecommunication, the Ministry of Internal Affairs and Communication ("MIC") has, not surprisingly, found itself to have a leading role to play in achieving such a futuristic society featuring AI and IoT. In fact, from as early as 2015 it has continuously held meetings of experts to address conceivable issues in light of the information and communication technology having reached an enhanced stage of development (details are given in II. below). Recently, the MIC's study group, now named "The Conference toward AI Network Society", published the Draft Artificial Intelligence Research and Development ("AI R&D") Guidelines for International Discussions and proposed a governance framework for the development of AI systems.

However, here lies a twist in policy formulation. The Science and Technology Basic Plan, which promulgated "Society 5.0," was adopted by the decision of the whole Cabinet, on the basis of the Report of the Council for Science, Technology and Innovation ("CSTI"). The latter Council is administered by the Cabinet Administration Office ("CAO") and not by the MIC. Furthermore, the study conducted by the MIC on enhanced information and communication technology preceded the Fifth Science and Technology Basic Plan. This implies a possible conflict between the MIC's study and the whole government's concept of "Society 5.0." It will thus be necessary to accommodate the outcome of the MIC's study, including the Draft AI Guidelines, into the CSTI's concept.

This article, in the following parts, first examines the proposed framework for the development of AI systems contained in the Draft AI R&D Guidelines (II.) and highlights the features of the Japanese approach towards the development of new technologies (III.). It then analyses the policy process concerning the Draft AI R&D Guidelines, including the process of conflict and accommodation mentioned above (IV.). Such an analysis will reveal the reality of the current policy formulation in Japan, which is different from what is widely believed. Some brief remarks about the role of law in relation to the emergence of new technologies concludes this article (V.).

---

1 See the Cabinet Administration Office's website on "Society 5.0" at *http://www8.cao.go.jp/cstp/english/society5_0/index.html*.

2 See the Cabinet Administration Office's website on "Society 5.0", *supra* note 1.

II. THE DRAFT AI R&D GUIDELINES

1. History of the Study Conducted by the MIC

As mentioned above, the MIC has organized a series of meetings to study the issues that arise when AI systems connected through a network come into actual use. The first of these meetings was held in early 2015 under the auspices of the "Study Group on the Vision of the Future in Relation to the Accelerating Intelligentization of ICT".[3] The members were twelve in number, consisting of engineers, economists and some industry representatives. No lawyers were involved in this Study Group. The Study Group concluded after several months with the publication of its Report in June 2015 and was succeeded by the "Conference for the Assessment and Evaluation of the Impact of the Intelligentization of ICT", which was soon renamed as the "Conference to Examine Networked AI".[4] The number of members was raised to forty-seven (including one advisor), and a few lawyers were now among them. As the Conference had become so large, three sub-groups were formed, these being responsible, respectively, for the economic, socio-human and legal-risk aspects. The Conference met for the first time in February 2016 and concluded its study with its Report in June of the same year.

"The Conference toward AI Network Society" was formed as the successor of the Conference of 2016, with thirty-four members and four advisors. Now that the Conference worked on promulgating R&D Principles, it formed the R&D Principles Sub-group as well as the Impact Assessment Sub-group. Having met from October 2016 to May 2017, the Conference published its 2017-Report and proposed the AI R&D Guidelines as its attachment.[5] The Conference was not dissolved this time but continued its discussions until it adopted the 2018 Report, which featured utilization Principles.[6] A few new

3 The materials and minutes of the meetings are available on MIC's website at *http://www.soumu.go.jp/main_sosiki/kenkyu/intelligent/index.html*. The original in Japanese begins with "interijento-*ka*", which literally translates as "intelligent-ization".

4 The materials and minutes of the meetings are available on MIC's website at *http://www.soumu.go.jp/main_sosiki/kenkyu/iict/index.html*.

5 THE CONFERENCE TOWARD AI NETWORK SOCIETY, *Draft AI R&D Guidelines for International Discussions* (28 July 2017), available at *http://www.soumu.go.jp/main_content/000507517.pdf* (hereinafter: AI R&D Guidelines). The Japanese version of the 2017 Report and its Appendices (including the Draft AI R&D Guidelines) are available at *http://www.soumu.go.jp/menu_news/s-news/01iicp01_02000067.html*. The materials and minutes of the meetings are available on MIC's website at *http://www.soumu.go.jp/main_sosiki/kenkyu/ai_network/index.html*.

6 The 2018 Report (in Japanese only) and its Appendices are available at *http:// www.soumu.go.jp/menu_news/s-news/01iicp01_02000072.html*. The materials and minutes of the meetings are available on MIC's website at *http://www.soumu.go.jp/main_sosiki/kenkyu/ai_network/index.html*.

members were added at this stage. Since the first Study Group of 2015, the meetings have been administered by the Institute for Information and Communications Policy, which is the policy research organization under the MIC.

2. *"Wisdom Network Society (WINS)" to be Achieved*

Through the meetings of the Conference, the concept of "WINS" has been formulated. WINS stands for "Wisdom Network Society" (in Japanese: 智連社会, *chiren shakai*), which is "a society where, as a result of the progress of AI networking, humans live in harmony with AI networks, and data/information/knowledge are freely and safely created, distributed, and linked to form a *wisdom network,* encouraging collaborations beyond space among people, things, and events in various fields and consequently enabling creative and vibrant developments."[7] Apparently, the concept emphasizes the positive aspect of networked AI systems and aims at the harmonious coexistence of human beings and AI systems. Interestingly, the focus is not on the data or information as such but on the "wisdom" to be established from the collection of data, information and knowledge.

The AI R&D Guidelines are prepared to achieve a WINS by facilitating the sound development of AI networks.[8] While the intention is not to regulate or control the development of AI systems, the policy orientation is clearly to introduce good governance for this emerging technology so that human society is not adversely affected by its unrestrained exploitation. One may find it to be an unusually strong commitment of the Japanese government to the good governance of technology.

The AI R&D Guidelines consist of four parts: Basic Philosophies, Definitions of Terms Used, Principles, and Comments on the Principles. Though the nine Principles, elaborated below, may attract the readers' attention because of their rule-like style, they are only a part of the Guidelines, and the latter must be read as a whole.

3. *Basic Philosophies of the AI R&D Guidelines*

In the part outlining the Basic Philosophies, the Guidelines introduce five such philosophies.[9] The first is "[t]o achieve *a human-centered society* where all human beings across the board enjoy the benefits from their life in harmony with AI networks, while human dignity and individual autonomy are respected" (emphasis in original). It is a reiterated commitment to the goal of protecting human society from the unwelcomed exploitation of AI technologies.

---

7 AI R&D Guidelines, *supra* note 5, 3, fn. 3 (emphasis in original).
8 AI R&D Guidelines, *supra* note 5, 3–4.
9 AI R&D Guidelines, *supra* note 5, 4–5.

The second Philosophy is the aspiration to lead the global discussion, while carefully avoiding regulation by hard rules. It reads "*[t]o share the Guidelines,* as non-binding soft law, *and their best practices internationally among stakeholders*" (emphasis in original). Any governance framework will surely be useless without international coordination, given the nature of the network (the Internet) as a medium transcending borders. Still, it is unusual for the Japanese government to reveal its aspiration to lead the global discussions; and it is still more unusual to take concrete steps towards that goal, as elaborated below.

The third Philosophy is the balancing of the "benefits and risks of AI networks". The Guidelines do not stop there but continue on to refer to "promot[ing] the benefits from AI networks through innovative and open R&D activities and fair competition" as well as "mitigat[ing] the risk that AI systems might infringe rights or interests, while fully respecting the value of the democratic society such as academic freedom and freedom of expression." The reference to these values is, again, a very strong commitment to the foundations of the modern civil society.

The fourth Philosophy is an acknowledgement of the importance of technological neutrality and the need to avoid placing "excessive burden" on developers. It is followed by the fifth Philosophy that the Guidelines will be subject to reviews and revisions as technology develops. Thus, the "Basic Philosophies" in fact comprise the two fundamental pillars – firstly, advocating human society and a commitment to the basic values of the civil society and, secondly, the intended nature of the Guidelines as a non-binding, technologically neutral and internationally shared instrument subject to future reviews.

4. Nine Principles for the Research and Development of AI Systems

On the basis of the above Philosophies, the Guidelines then introduce nine Principles addressed to developers engaged in the research and development of AI systems. Of the nine Principles, one is for the sake of the sound development of AI networking and the promotion of the benefits of AI systems, six address the mitigation of risks associated with AI systems and the remaining two are related to the acceptability of AI systems among users.[10]

The first principle is the Principle of Collaboration, which requires that "[d]evelopers should pay attention to the interconnectivity and interoperability of AI systems." The comments on this Principle elaborate that the developers should make efforts towards sharing information that is effective in ensuring interconnectivity and interoperability, conformity with any international standards, standardization of data formats and openness of interfaces and

10 AI R&D Guidelines, *supra* note 5, 7–8.

protocols. The comments further note that developers need to be aware of unintended risks resulting from interconnection or interoperations of its own AI system with other AI systems and require the developers to make efforts to advance open and fair treatment concerning the licensing of intellectual property, including standard essential patents, which will be useful for interconnectivity and interoperability.

The following six Principles, relating primarily to the mitigation of risks, are the Principles of Transparency, Controllability, Safety, Security, Privacy and Ethics. Among them, the Principle of Transparency requires developers to "pay attention to the verifiability of inputs/outputs of AI systems and the explainability of their judgments."[11] This will be most problematic with self-learning AI systems. The Principle does not excuse a developer if the self-learning AI has grown to cause harm to the user or to society after the developer has finished coding the initial system, and it demands that a developer enable an *ex post* verification of how the AI system learned by itself (and ended up becoming harm-inflicting).

The third Principle, the Principle of Controllability, encourages developers to "conduct verification and validation in advance" of applying the AI system in actual life.[12] The comments on this Principle also mention that human supervision and counter-measures, such as shutting down the AI system or cutting it off from the network, should be considered.

The next principle, the Principle of Safety, suggests that developers see to it that "AI systems will not harm the life, body, or property of users or third parties through actuators or other devices." The comments on this Principle encourage developers to conduct verification and validation in advance, implement preventive measures useful for intrinsic and functional safety[13] and explain to stakeholders the design of the AI system as regards the priority given to the personal life, body and property involved. The comments specifically mention the benefit of referring to international standards, though it is not entirely clear which international standards are meant.

The fifth Principle is the Principle of Security, which is tautologically elaborated as "[d]evelopers should pay attention to the security of AI systems." The comments provide, first, that international guidelines such as the

11 "Explainability" is the term used in the English version of the Guidelines (AI R&D Guidelines, *supra* note 5, 7). It is obviously an artificial and strange word. In Japanese, *setsumei sekinin* (説明責任) does not sound strange, though the word is usually understood to mean "accountability".

12 AI R&D Guidelines, *supra* note 5, 9.

13 According to the comments on the Principle of Safety, "intrinsic safety" means the "reduction of essential risk factors such as kinetic energy of actuators", while "functional safety" means the "mitigation of risks by operation of additional control devices such as automatic braking". See AI R&D Guidelines, *supra* note 5, 10.

OECD Guidelines for the Security of Information Systems and Networks[14] must be respected. The comments note that, in the context of AI systems, the security of information includes reliability (whether the operations are performed as intended and not steered by unauthorized third parties) and robustness (tolerance to physical attacks and accidents) of AI systems, besides the confidentiality, integrity and availability that are generally required. Then, as in the previous Principles, developers are encouraged to consider conducting verification and validation in advance and to "take measures to maintain the security to the extent possible", which intends to implement the concept of "security by design."

The Principles of Safety and Security are followed by the Principle of Privacy. This Principle demands that developers "take it into consideration that AI systems will not infringe the privacy of users or third parties." The comments elaborate that privacy includes "spatial privacy" (peace of personal life), "informational privacy" (protection of personal data) and the "secrecy of communications".[15] The comments then require developers to "evaluate the risks of privacy infringement and conduct [a] privacy impact assessment in advance" and to "take necessary measures […] to avoid infringement of privacy at the time of the utilization." It is noted that the latter requirement means "privacy by design". As in the case of the Principle of security, reference is made to the OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data.[16]

The final Principle that belongs to the category of mitigation of risks associated with AI systems is the Principle of Ethics. Here, developers are required to "respect human dignity and individual autonomy in R&D of AI systems." As elaborated in the comments, developers are encouraged to "pay particularly due [sic] consideration to respecting human dignity and individual autonomy, in light of discussions on bioethics, etc" and to "take necessary measures so as not to cause unfair discrimination resulting from prejudice included in the learning data of the AI systems." Furthermore, developers are advised to "take precautions to ensure that AI systems do not unduly infringe the value of humanity." In the context of the last part, International Human Rights Law and the International Humanitarian Law (capitalized in original) are referred to.

The eighth and ninth Principles are for the sake of acceptability of AI systems among its users. On the one hand, the Principle of User Assistance requires developers to ensure that AI systems are supportive of users and that

14 Available at *http://www.oecd.org/sti/ieconomy/15582260.pdf*.

15 Secrecy of communications is enshrined in Art. 21(2) Japanese Constitution (*Nihon koku kenpō* of 1946), as well as in Art. 4(1) Telecommunications Business Act (*Denki tsūshin jigyō-hō,* Law No. 86/1984).

16 Available at *http://www.oecd.org/sti/ieconomy/oecd_privacy_framework.pdf*.

users are given "opportunities for choice in appropriate manners." According to the comments, this requirement includes making available such interfaces that are easy to use, enabling users to make timely and appropriate choices, and having regard for the use of AI systems by socially vulnerable people through, for example, a universal design.

On the other hand, there is the Principle of Accountability, which requires developers to "fulfill their accountability to stakeholders including AI systems' users." The comments divide this requirement into two: one is to "provide users with the information that can help their choice and utilization of AI systems" while the other is to provide users with "information and explanations about the technical characteristics of the AI systems they have developed" as well as to ensure "active involvement of stakeholders" through dialogue.

III. The Approach of the Draft AI R&D Guidelines

1. AI R&D Guidelines as State-induced Self-regulation

In terms of their form, the Draft AI R&D Guidelines are not intended as a draft of a legally binding instrument. In this sense, it contrasts with, for example, the motion to introduce a set of Robo-laws by a group of European Parliament members.[17] The Guidelines emphasize the benefit of "non-regulatory and non-binding soft law".[18] Although there is no explicit reason given as to why a legally binding instrument is found to be inappropriate, the reference to the rapid development of R&D and the use of AI systems in the second Philosophy may reveal the thoughts behind this preference.

When a new technology that may affect people's lives emerges, developing a norm to prevent harmful exploitation of the technology is useful for acquiring the public's consent to such a technology. In some European jurisdictions, a strict liability regime was occasionally employed for that purpose.[19] Although Japanese private law has been influenced by German law in many respects, such an approach has not become common in Japan.[20] The strict liabil-

17 Committee on Legal Affairs, Draft Report with recommendations to the Commission on Civil Law Rules on Robotics (2015/2103 (INL)), available at *http://www.europarl.europa.eu/oeil/popups/ficheprocedure.do?lang=en&reference=2015/2103(INL)#documentGateway*.

18 AI R&D Guidelines, *supra* note 5, 3.

19 See C. von Bar (ed.), Principles of European Law / Non-Contractual Liability Arising out of Damage Caused to Another (Oxford 2009) 724–735 and 738–741.

20 See S. Kozuka, Strict Liability and State Indemnification under Japanese Law: The New Space Activities Act Compared with the scheme on Compensation for Nuclear Damages, ZJapanR / J.Japan.L. 43 (2017) 3.

ity for harm resulting from mining under the Mining Act,[21] as well as under the post-war Act on Compensation for Nuclear Damages,[22] are rare exceptions. Later, legislation introducing strict liability was either a response to already widespread pollution damage in the 1970s or, in the case of the most recent Space Activities Act,[23] more focused on the government's indemnification of operators for the purpose of enhancing their competitiveness in the global market. It is, therefore, not surprising that the initiatives concerning the norms on AI systems do not intend to introduce strict liability rules.

In fact, the preference for a non-binding, non-legal instrument is not exceptional with regard to norms on AI systems. The Asilomer Principles, one of the most famous statements establishing a set of norms on AI systems, also adopt this approach.[24] However, the Asilomer Principles were drafted by academics and industry leaders in this sector and may be regarded as being of a purely private nature. Contrary to such an initiative, Japan's AI R&D Guidelines have been drafted under the auspices of the Japanese government. In this respect, it can be regarded as a state-induced self-regulation.

Elsewhere, this author has analysed state-induced self-regulation in Japan.[25] In one case, the relevant statute delegates a self-regulatory power to elaborate rules. In another case, self-regulation is used to have industry members commit to a certain interpretation of a statutory provision. There are, furthermore, cases where state-induced self-regulation is relied on when there is no statutory regulation yet. Apparently, the rules on AI systems belong to this last group. Where the introduction of a new law has been found to be not preferable for fear that early regulation could stifle the developing technology, state-induced self-regulation has been considered appropriate to generate public trust in the emerging use of AI systems.

2. The Absence of Legal and Liability Issues

The substantive norms included in the AI R&D Guidelines reflect the features and form of the Guidelines. First, the liability issue is carefully avoided, with the term "risk" being mentioned several times. This may connote that legal liability is considered as one of the "risks" that developers of AI system must be aware of. Still, a question about legal consequences may arise, as is the

21 *Kōgyō-hō,* Law No. 289/1950.

22 *Genshi-ryoku songai no baishō ni kansuru hōritsu,* Law No. 147/1961.

23 *Jinkō eisei no uchi'age oyobi jinkō eisei no kanri ni kansuru hōritsu* [Act Concerning the Launch and Control of Man-made Satellites], Law No. 76/2016.

24 FUTURE OF LIFE INSTITUTE, Asilomar AI Principles, available at *https://futureoflife.org/ai-principles/.*

25 S. KOZUKA, Self-regulation Induced by the State in Japan, in: Baum/Bälz/Dernauer (eds.), Self-regulation in Private Law in Japan and Germany (Cologne 2018) 109.

case, for example, in respect of the Principles of Controllability and Safety. Both Principles call for assessment, verification and validation in advance by developers. A lawyer may see possible liability when such a process is neglected. The same may be said of the Principle of User Assistance. Apparently, the drafters of the Guidelines do not wish to prejudice legal standards for the establishment of liability for instances where an incident occurs in the course of the use of an AI system.

Secondly, a somewhat related feature of the Guidelines is that the enforcement issue is nowhere mentioned. In the Principles of Safety, Security and Privacy, the values recognized in each Principle are to be ensured "by design." It is indeed a widely accepted view that the technical design of an AI system will be more effective in protecting these values than any legal rules. On the other hand, how such "by design" requirements are to be enforced, and what the outcome should be if a developer fails to incorporate these requirements into the design of an AI system, is seldom discussed. The AI R&D Guidelines are also not helpful in this regard.

Thirdly, the Guidelines appear to consider that an AI system should not be a "black box" for the eyes of the users. With its use of the strange expression "explainability," the Principle of Transparency requires that a judgment by an AI system be "explained". The Principle of Accountability similarly emphasizes the need to provide users with information and explanation. Probably with good intentions, such emphasis on explanation could imply that the developer is responsible only for explaining the mechanism, while the user assumes all the risks for the outcome of using the AI system, presumably with knowledge of these risks. However, the concept of accountability is not the same as (solely) a duty to explain.[26] Depending on the circumstances, the provider of an AI system may be required to assume responsibility for the results from its use, whether or not the explanation as to how the system works has been given. It is thus obvious that the Guidelines' emphasis on explanation derives from its preference not to prejudge any legal responsibility.

On the one hand, these features of the AI R&D Guidelines seem to be consistent with the choice of non-binding, non-legal self-regulation. On the other hand, if this is the intention, the question may arise why a state-induced instrument and not a purely private self-regulatory form was preferred. A clue to answering this question may be found in the fact that all the Principles are drafted as messages to developers. In other words, the aim of introducing the Guidelines – including among other elements the Principles – may be that developers will be justified in relying on them as a "safe harbour." If not a complete exemption from liability in the legal sense, compliance with the

26 B. A. GARNER (ed.), Black's Law Dictionary (10th ed., St. Paul/Minnesota 2014) defines the term "accountable" as "responsible; answerable."

Guidelines may nevertheless establish that developers are not operating under a bad motive.

3. The Aspiration to Lead the Global Discussions

Furthermore, because AI systems, especially when connected to a network, will be used globally, the safe harbour should be common internationally. Indeed, in April 2016, on the occasion of the G7 summit in Japan, the then Minister for Internal Affairs and Communications, Sanae Taka'ichi, introduced the draft of the AI R&D Guidelines and proposed that international debate on it take place in an appropriate forum, such as the OECD.[27] In response, the OECD co-sponsored an international symposium with Japan's MIC on AI policies in October 2017.[28] Two members of the Conference toward AI Network Society participated and presented the Draft AI R&D Guidelines. Given that Japan has been rather reactive and selective in respect of global efforts towards the unification and harmonization of laws,[29] actively taking the lead in the global debate over AI policies is exceptional.

In the international context, one might expect a cultural hue in the Japanese Guidelines. This may be all the more the case as it is often said that the love of robots is widespread in Japanese society.[30] However, no trace of such cultural aspects is to be found in the Guidelines. Whether or not such an emotional uniqueness really exists in the Japanese, the drafters thought that it has no relevance to the norms for the research and development of AI systems. After all, if there is an aspiration to lead the global debate on the subject, an emphasis on cultural aspects would be useless if not a hindrance.

IV. THE PROCESS OF FORMULATING THE GUIDELINES

The process of drafting the AI R&D Guidelines has allowed experts as well as industry and consumer representatives to come together and discuss the vari-

---

27 See *Follow up Report of the Charter and the Joint Declaration from the 2016 G7 ICT Ministers' Meeting* by Japan as 2016 G7 Presidency (2016) 3, available at *http://www.soumu.go.jp/joho_kokusai/g7ict/english/pdf/113007_02.pdf*.

28 *http://www.oecd.org/going-digital/ai-intelligent-machines-smart-policies/*.

29 See H. SONO, Going Forward with Uniform Private Law Treaties: A Study in Japan's Behavioral Pattern, Japanese Yearbook of International Law 60 (Tōkyō 2017) 10.

30 The "love" of robots by the Japanese is widely recognized and is often accompanied by cultural explanations. See, for example, M. PROSSER, Soon We'll All Love Robots the Way Japan Loves Robots, *https://singularityhub.com/2016/10/20/soon-well-all-love-robots-the-way-japan-loves-robots/#sm.0001f8b356my4elptz912pa3x6ic6;* J. ITO, Why Westerners Fear Robots and the Japanese Do Not, *https://www.wired.com/story/ideas-joi-ito-robot-overlords/*. Needless to say, whether this is only a myth or indeed the truth requires careful examination.

ous issues at meetings organized by the MIC. Given that it is common knowledge that the making of law or policy in Japan takes place at councils *(shingi-kai)* of government agencies,[31] this may sound like a further instance of such system in practice. However, the reality is more complex than it may appear.

Institutionally, councils are only those that are prescribed by the Laws on the Constitutions of each ministry or by Article 8 of the National Government Organization Act.[32] Such councils have the mandate of responding to inquiries by the minister in question. In the case of the MIC, Article 8 of the Law on the Constitution of the MIC provides for the Local Public Finance Council, the Administrative Complaint Review Board, the Information Disclosure and Personal Information Protection Review Board, the Supervisory Commission for Public-Private and Private-Private Competitive Tenderings, the Committee on the System of Evaluating Incorporated Administrative Agencies, the Central and Local Government Dispute Management Council, the Telecommunications Dispute Settlement Commission, the Radio Regulatory Council and the Statistics Commission. Furthermore, pursuant to Article 8 of the National Government Organization Act, there is also the Council of Information and Communications under the Cabinet Ordinance.[33] It is customary that amendments to important laws under the jurisdiction of the MIC or the adoption of important policies are examined by these Councils.

In some cases, however, discussions take place not in the councils but in ad hoc "study group" meetings. These study groups are not founded on a law but by a councillor of the ministry (usually ranked second to the deputy secretary of the ministry), the head of a bureau or even the head of a division, as the case may be. They are useful when focused discussions on a particular topic are required or when a series of concentrated discussions are needed. However, the Conference toward AI Network Society, which promulgated the Guidelines and published the Report including them, is different from such an ad hoc study group in that its meetings (though physically taking place in the building of the MIC) are not administered by a section of the MIC but by its policy research institute. It indicates the somewhat informal nature of the Conference.

This informal nature of the Conference is, on the one hand, the outcome of choosing a non-legal form. Had it been a proposal for law reform, the fora of discussions could not have been anything other than a formal council, or at

31 E. HARARI, Resolving and Managing Policy Conflict: Advisory Bodies, in: Eisenstadt/Ben-Ari (eds.), Japanese Models of Conflict Resolution (London 1990); G. NOBLE, Reform and continuity in Japan's *shingikai* deliberation councils, in: Amyx/Drysdale (eds.), Japanese Governance: Beyond Japan Inc. (London 2003) 113.

32 *Kokka gyōsei soshiki-hō,* Law No. 120/1948.

33 *Jōhō tsūshin shingi-kai-rei* [Ordinance on the Council of Information and Communications], Cabinet Ordinance No. 271/2000.

least an ad hoc study group. On the other hand, the informal nature of the fora may have resulted in the instrument being of a more theoretical rather than a narrow policy-oriented nature. As elaborated above, the Conference's Report starts from the concept of WINS, on which both the Principles and the founding Philosophies are based. WINS is itself a picture of the society that the Conference considers desirable when networked AI is fully developed. It is too abstract, and maybe too idealistic, as to be an actual policy goal, and the Report does not advocate it as such. Still, the fact that the Guidelines are based on the concept of WINS means that the instrument is not a mere product of compromise among interested parties, but that it is fuelled by a vision about the future society. The reference to the foundational values of civil society in the five Philosophies may also derive from such an idealistic nature of the Guidelines.

Another outcome of employing an informal forum is that both the Report and the Guidelines are not limited to the subjects under the jurisdiction of the MIC. More specifically, the Report covers the use of AI for autonomous vehicles and medical services,[34] which obviously overlaps with the responsibilities of other governmental agencies. The fact that the Conference was administered by the policy research institute may have contributed to the scope of the Report not being limited to the strictly defined jurisdiction of the MIC.

Quite naturally, though, the instrument's broad scope has led to conflicts with activities of other ministries subsequent to the publication of the Guidelines. For example, as regards autonomous vehicles, the Strategic Headquarters for the Promotion of an Advanced Information and Telecommunications Network Society ("IT Strategic Headquarters"), having set up a sub-group, adopted the Outline of Institutional Reform Concerning Autonomous Vehicles in April 2018.[35] While this Outline mainly addresses regulatory aspects, such as safety guidelines and traffic rules, the civil liability aspect was discussed in another forum – namely, the Study Group on the Liability Arising from Autonomous Vehicles, established by the Ministry of Land, Infrastructure and Transport. The Study Group, as one of the ad hoc study groups, published its report in March 2018, concluding that some modifications to, but not a radical restructuring of, the existing liability scheme under the Act on Liability and Compensation for Damages from Automobiles were necessary.[36] As regards the use of AI systems for medical services, the Ministry of

34 See 2017-Report, *supra* note 5, 35, and Appendix 3 on the case studies regarding the use of AI systems.

35 KŌDO JŌHŌ TSŪSHIN NETTOWĀKU SHAKAI SUISHIN SENRYAKU HONBU, *Jidō unten ni kakaru seido seibi taikō* [Outline of Institutional Reforms Concerning Autonomous Vehicles] (17 April 2018), available (only in Japanese) at *https://www.kantei.go.jp/jp/singi/it2/kettei/pdf/20180413/auto_drive.pdf*.

36 KOKUDO KŌTSŪ-SHŌ JIDŌ-SHA-KYOKU, *Jidō unten ni okeru songai baishō sekinin ni kansuru kenkyū-kai hōkuku-sho* [Report of the Study Group on Liability for Damages

Welfare, Health and Labour set up another ad hoc study group, the Forum for the Promotion of AI in the Health and Medical Sector, which concluded its debates and published a report in June 2017.[37]

Under the traditional Japanese political system, which one commentator termed "compartmentalised pluralism", intruding into the territory of another government agency has been considered as indecent for bureaucrats.[38] However, as an overlap of deliberations is the norm rather than the exception in the modern, complex society, the Cabinet Secretariat has increased its power of coordination in the last couple of decades.[39] In recent years, any policy demanding a budgetary commitment must be entered into the instrument adopted by the Cabinet in the first half of that year, a document generally known as the Growth Strategy.[40]

Although the AI R&D Guidelines do not require budgetary expenses, there is obviously a need to coordinate the Guidelines with the "Society 5.0" concept, which is the concept adopted by the Cabinet and committed to in the Science and Technology Basic Plan. In fact, the Cabinet Administration Office, which administers the Conference for the Strategy of AI Technology, set up the Study Group on the Principles for a Human-Centred AI Society in April 2018.[41] Apparently, the Principles that this Study Group intends to adopt are similar to the AI R&D Guidelines, and the latter have been submitted to the Study Group as material for discussion. Therefore, it is likely that the outcome will be a more or less similar set of rules. Still, it is yet to be seen whether this Study Group, a more formal forum than the Conference, makes any significant modifications to the AI R&D Guidelines. What may be more intriguing is

---

concerning Autonomous Vehicles] (March 2018), available (only in Japanese) at *http://www.mlit.go.jp/common/001226365.pdf*.

37 HOKEN IRYŌ BUN'YA NI OKERU AI KATSUYŌ SUISHIN KONDAN-KAI, *Hoken iryō bun'ya ni okeru AI katsuyō suishin kondan-kai hōkoku-sho* [Report of the Forum for the Promotion of AI in the Health and Medical Sector] (27 June 2017), available (only in Japanese) at *https://www.mhlw.go.jp/file/05-Shingikai-10601000-Daijinkanbou kouseikagakuka-Kouseikagakuka/0000169230.pdf*.

38 The concept of "compartmentalized pluralism" *(shikirareta tagen shugi)* was advanced by S. SATŌ/T. MATSUZAKI, *Jimin-tō seiken* [The LDP Government] (Tōkyō 1986). For a critical review of this and other related concepts see J. C. CAMPBELL/E. SCHNEIDER, Fragmentation and Power: Reconceptualizing Policy Making under Japan's 196\55 System, Journal of Political Science 9 No. 1 (2008) 89.

39 T. SHINODA, Japan's Cabinet Secretariat and its Emergence as Core Executive, Asian Survey 45 No. 5 (2005) 800–821.

40 From 2013 to 2016, each year's instrument had formally been titled the "Japan Revitalization Strategy"; since 2017 it has been named the "Future Society Investment Strategy".

41 The materials and minutes of the meetings are available on the Cabinet Administration Office's website at *http://www8.cao.go.jp/cstp/tyousakai/humanai/index.html*.

to see whether this Study Group maintains the WINS concept and continues to advocate in favour of civil society values as the founding philosophies of the AI Guidelines.

V. Conclusions

The AI R&D Guidelines give insights into how Japan is responding to the development of new technology and what role the law plays in these developments. While the existence of legal risks is recognized and the need to ensure the public's acceptance of the technology is considered essential, Japan has preferred not to introduce a new legal framework, still less liability rules, and has instead chosen to promulgate non-binding "Principles." It is also intriguing to see that Japan is not hesitating to take initiatives in order to disseminate these Principles to other parts of the world in cases where fragmented regulation seems problematic to the industry in question. Such international aspirations notwithstanding, domestic politicking still matters within Japan. In particular, the power of the Cabinet Secretariat to coordinate has become so significant that no government agency can ignore it. It remains to be seen how successful the Japanese initiatives to introduce the AI R&D Guidelines as the governing framework for this new technology will be in the end.

Summary

The article considers the "Draft Artificial Intelligence Research and Development Guidelines for International Discussions" (hereinafter: AI R&D Guidelines), published in 2017 by a study group of the Ministry of Internal Affairs and Communication, namely by the Conference toward AI Network Society. It begins by giving an overview of the development process and then analyses the content of these Guidelines, in particular its nine Principles. While these are formulated in an almost prescriptive tone, they are actually intended as non-binding guidelines for the development of AI systems. This aim is reflected in the five Basic Philosophies, which precede the Principles in the Guidelines. The Principles are addressed to developers of AI systems and cover various aspects, as indicated by their names: collaboration, transparency, controllability, safety, security, privacy, ethics, user assistance and accountability. As the commentary on these Principles reveals, the Guidelines propose a framework that ensures compatibility of AI systems, mitigates risks associated with AI, increases the acceptance of AI by its users and holds developers accountable to stakeholders.

Based on this information, the article addresses the question of why the form of non-binding, state-induced self-regulation was chosen over legally binding norms. This seems to be linked strongly to the Japanese approach towards the development of new technologies, according to which a strict liability regime is usu-

ally not adopted. Indeed, the AI R&D Guidelines follow this approach. Conversely, the Japanese government has been unusually proactive in the international discussion of AI systems. The article's depiction of policy-making in Japan, where discussions are often held by study groups established by a ministry councillor rather than under a legal measure, further helps to explain the choice for the form of the Guidelines. The extent to which Japan's initiative is successful remains to be seen.

(The Editors)

ZUSAMMENFASSUNG

Der Beitrag befasst sich mit den 2017 von dem Ministerium für Interne Angelegenheiten und Kommunikation bzw. der Konferenz für eine durch künstliche Intelligenz vernetzte Gesellschaft veröffentlichten „Draft Artificial Intelligence Research and Development Guidelines for International Discussions" (nachfolgend: AI R&D Richtlinien). Der Verfasser gibt zunächst einen Überblick über die Entstehung der Richtlinien und geht sodann auf ihren Inhalt, insbesondere die neun „Principles" ein. Während diese im Ton beinahe als verbindlich klingen, sind sie jedoch als nicht bindende Richtlinien für die Entwicklung von AI gedacht. Dieses Bestreben spiegelt sich in den fünf den Prinzipien vorausgehenden „Basic Philosophies" wider. Die Prinzipien richten sich an Entwickler von AI Systemen und behandeln verschiedene Aspekte. Diese lassen sich in ihren Bezeichnungen erkennen: Zusammenarbeit, Transparenz, Kontrollierbarkeit, Sicherheit, Privatsphäre, Ethik, Benutzerunterstützung und Verantwortlichkeit. Die Kommentare zu den Prinzipien verdeutlichen, dass die Richtlinien ein System vorschlagen, in dem die Kompatibilität von AI Systemen sichergestellt, die mit AI verbundenen Risiken minimiert, die Akzeptanz der Nutzer von AI erhöht, und die Entwickler gegenüber den Nutzern und anderen Beteiligten verantwortlich gemacht werden.

Auf dieser Informationsgrundlage geht der Beitrag dann der Frage nach, weshalb die Regelsetzung mit den Richtlinien in Form von nicht bindender, durch den Staat initiierter Selbstregulierung erfolgt ist und keine rechtlich bindenden Normen geschaffen wurden. Der Grund hierfür liegt wohl einerseits in dem japanischen Ansatz zur Regulierung von neuen Technologien, die normalerweise nicht auf das Regime der verschuldensunabhängigen Haftung zurückgreifen. Dies ist auch der Ansatz der AI R&D Richtlinien. Auf der anderen Seite ist die Mitwirkung der japanischen Regierung in der internationalen Diskussion um AI Systeme ungewöhnlich proaktiv. Ein weiterer Grund könnte in der Politikgestaltung Japans liegen, wie beispielsweise der Tatsache, dass Diskussionen oft in Arbeitsausschüssen („study groups") geführt werden, die jedoch nicht per Gesetz, sondern aufgrund der Entscheidung eines Ministerrats eingesetzt werden. Es gilt abzuwarten, in welchem Umfang die Initiative Japans Erfolg hat.

(Die Redaktion)