Good Data
Data-sharing, public trust and willingness

BRIEFING PAPER
APRIL 2021

WORLD ECONOMIC FORUM

A new social contract on data use

COVID-19 has revealed social disparities and delays in digitalization in many countries, exposing the large gap between current reality and the ideal digital society – the inclusive, seamless vision known in Japan as Society 5.0[1].

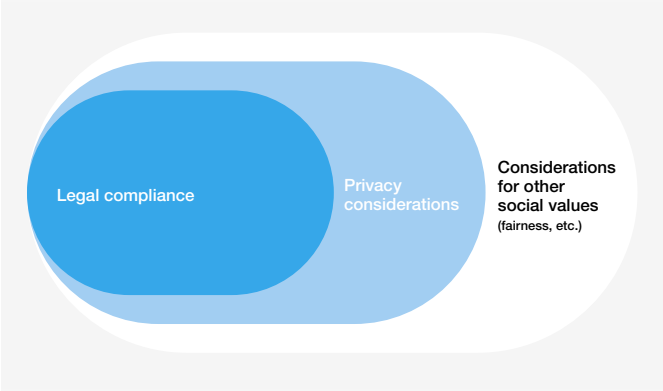
The theme of the World Economic Forum Annual Meeting 2021 is the "Great Reset"[2], advocating building the foundations of economic and social systems for a future that is just, sustainable and resilient. Instead of choosing between public health and the economy, for instance, we need a brand new social contract that ensures sustainability while protecting human life, dignity and social justice.

Data can be a powerful tool in this effort. Data is not only a "primary" asset – with value to the organizations that collect and use it for specific purposes – but also a "secondary" one, in which additional value is created through broader circulation. How to safely promote this secondary use of data is a major topic of discussion around the world.

Data-sharing issues

It would be unreasonable to expect the secondary data use to have only a bright side. The key to "good data" lies in how it is handled: with consideration for privacy, fairness, social justice, and ethics, not to mention legal compliance. Data-governance systems need to address aspects of data handling such as Ethical, Legal and Social Implications (ELSI), Ethical, Legal, Social and Economic (ELSE) impact, and Responsible Research and Innovation (RRI). By doing so, they can gain the trust of the public, as individuals and as data subjects, and increase people's willingness to provide data.

FIGURE 1 **Scope of considerations which should be given in data utilization**



What sort of data use promotes trust and willingness? Considerations about ethics and social values are difficult to articulate, and values and cultures are diverse. It can be unclear how much consideration should be given to which factors in order to support "good data" (or to avoid damaging social confidence). Several global regulations on such issues have been introduced, such as the eight OECD core privacy principles[3], the Asilomar AI Principles[4], the OECD's Recommendation of the Council on Artificial Intelligence[5], and the Ethics Guidelines for Trustworthy Artificial Intelligence (AI) by the EU[6].

However, when we turn to how data is handled in real-world settings, these principles and guidelines may not function sufficiently, as they do not indicate specific interpretations and methods for practical application. More useful, often, are example cases in medical research and advanced medical treatment. In these areas, governance systems have been established to balance ethical considerations and innovation (medical progress), as can be seen in ethical principles such as the Declaration of Helsinki[7]. However, as innovators continue to enter the healthcare market – including start-ups and companies from other industries – familiarity with medical research and its historical background is far from universal. Therefore, society needs initiatives to balance ethical considerations with innovation in data-sharing in the private sector.

Limits of the notice and consent model

In order to gain trust in data utilization and strengthen people's willingness to provide data, the data governance model known as notice and consent, in which each individual data subject is notified and asked for consent[8] each time his or her data is collected, must be revisited. In reality, only a few users read privacy policies carefully, a situation that is likely to be exacerbated by the development of the internet of things (IoT)[9]. In addition, it can be difficult to apply the notice and consent model to elderly people whose cognitive functions have declined because the data subject's ability to understand his or her situation and to reason – the prerequisite of notification and consent – is compromised. Thus, a new scheme for data-handling that can be trusted by cognitively impaired elderly people may need to be developed.

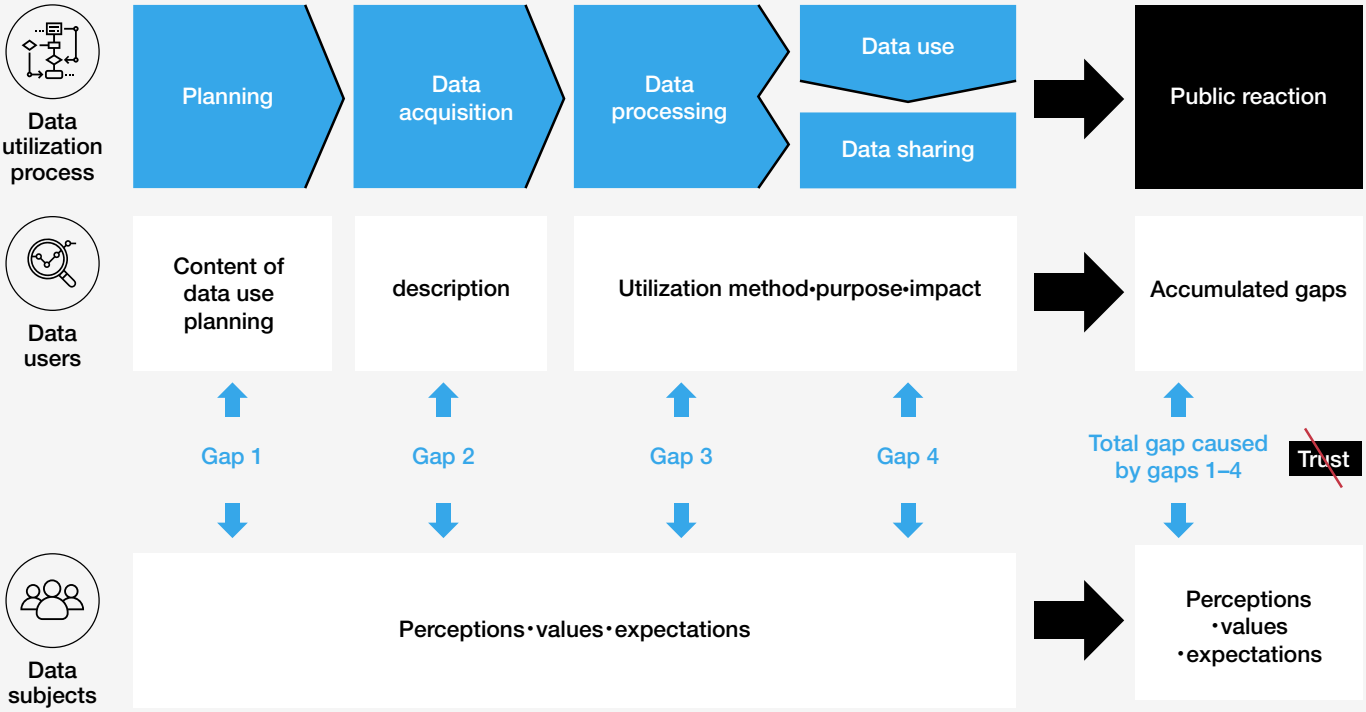
Trust and willingness framework

Framework of public trust

Some studies on trust have shown that data subjects will trust data users if the motives of the data users are consistent with individuals' perceptions, values and expectations. Matching the perceptions, values and expectations of data subjects and the public with the purposes and motives of data users fosters data subjects' willingness to provide data. However, gaps can occur between the methods and impacts of data utilization and the perceptions, values and expectations of individuals and the public at the data utilization phase. This gap can cause "negative surprises" to individuals and the public, resulting in broken and lost trust.

In each phase of data utilization, why and what kind of gaps arise between the perceptions, values and expectations of individuals as data subjects, and actual data use by data users, along with its impacts, are outlined in Figure 2.

FIGURE 2 **Gaps created in data utilization**



The following is a description of gaps **1** to **4** that occur in each phase of data utilization shown in Figure 2.

BOX 1 **Gaps created in data utilization**

Gap 1: Gap arising from uncertainty about methods used to deal with data in an ethically appropriate manner at the planning phase

Gap 2: Gap arising from the way explainers communicate with data subjects and the data subjects' literacy at the data acquisition phase

Gap 3: Gap arising from differences from the planning stage or unforeseen effects during the implementation stage, including data processing, utilization, and third-party provision

Gap 4: Gap arising from biased reporting and information, excessive simplification and distortion by the media and SNS, etc.

In light of these gaps, the following two points should be considered by data users to gain trust from the public in data utilization.

1. Sustainable approaches to reduce the gaps that lead to losses and breaks in public trust in data utilization

2. Initiatives to build multistakeholder trust through the use of Fourth Industrial Revolution technologies

Good data framework

Ethically sound data-handling framework

As a framework for narrowing gaps that can lead to the collapse of society's trust in the use of data, we have divided the data-handling process into five phases (planning, data acquisition, data processing, implementation and data provision) and presented the points that data users should keep in mind in each phase. This is intended to serve as a practical framework for data users.

1 Planning phase

1.1 Appropriateness of data use

In the planning phase, the values and benefits to be realized by data utilization should be clearly determined.

- In some cases, it may be more effective to use tools other than data, for reasons such as operational stability, economic rationality and number of samples.
- The interests of those who plan to use data (e.g., recruiters) do not necessarily coincide with those of data subjects (e.g., prospective employees).

1.2 Purpose of data use

The type of data involved and the purpose for which it is to be used should be clarified.

- It should be clearly stated whether profiling will be conducted and, if so, details should be disclosed.
- Regardless of whether profiling is conducted, the purpose of data use should be clearly indicated to data subjects at the time of data acquisition.

1.3 Ethics by Design in organizations using data

Service design to reduce legal, ethical and social risks should be considered.

- The IEEE, an academic organization, published a report titled, Ethically Aligned Design,[10] and is aiming to develop a standard (P7000 series).

1.4 Whether it is possible to obtain understanding and support from data subjects and the public regarding the purpose of data use and the business model

Taking into consideration the perspectives and backgrounds of the public, it should be examined and confirmed whether it is possible to gain understanding and support for the purpose of data use and the business model without causing anxiety, concern or discomfort.

- Science communication efforts such as science cafés in the UK and consensus conferences in Denmark are an effective approach.
- In the Legitimate Interest Assessment published by the European Commission, the importance of capturing the perspectives of citizens to balance the interests of individuals and those of companies is indicated, and it can be carried out by asking yourself questions such as, "Are you happy to explain it to them?"; "Are some people likely to object or find it intrusive?"

2 Data acquisition phase

2.1 Appropriate means of acquiring data and obtaining consent

Data should be acquired through appropriate means. Also, when acquiring personal information, consent should be obtained through appropriate methods.

- The GDPR states that users should be free to choose and should not be forced to give consent. Opportunities to consent for the use of services, participation in clinical trials and use of data should be made available independently.

2.2 User interfaces and other components for data acquisition

Methods to effectively support the decision-making of data subjects should be devised.

- There is a view that the text of privacy policies is difficult and lengthy and thus exceeds human cognitive ability.
- As this is the only touchpoint with the data subject, it is desirable that appropriate care be taken.

2.3 Bias in datasets

Bias in datasets should be examined.

- A data set is only a slice of an event or phenomenon (of the physical world) and there is a risk that bias will lead to underrepresenting or overrepresenting certain communities.
- If bias or discrimination against race or gender is present in the real-world data itself, it may further contribute to that bias or discrimination.

3 Data processing phase

3.1 Privacy infringement risks in data analysis and evaluation

Risk of privacy infringement in data analysis and evaluation should be considered.

- The higher the accuracy of the prediction, the more likely it is to be perceived as truthful by third parties, and thus coordination with the legal interests of the data subjects is required. At present, there are no specific guidelines on the balance between the legal interests of data users and those of data subjects, and public discussion and consensus building are being sought.

– For the time being, it is important to consider privacy-infringement risks and seek appropriate disciplines.

3.2 Risks to stakeholders other than data subjects resulting from data processing, analysis and evaluation

Risks of and effects on stakeholders other than data subjects in data analysis and evaluation should be considered.

– For example, in the case of genetic data, it is desirable to consider the impact on group privacy, since blood relatives share a portion of the same genes.

– In addition, when data is processed by someone other than the data user, responsibility may fall on the data processor. An example of such a case is when data from a medical institution is processed anonymously by the medical institution and then provided externally as non-personal information.

3.3 Introduction of fair data analysis and learning technologies

Consideration should be given to the introduction of technological measures that take into account social justice and anti-discrimination in light of the product's purpose.

– In order to address prejudices and biases that may be latent in datasets, it is recommended that measures be taken to remove sensitive features and elements that lead to discrimination or to reduce the influence of them.

– An ethical problem that can arise is: "Should we achieve equality and fairness at the expense of algorithmic prediction and model accuracy?" A multifaceted approach should be considered, such as addressing values that cannot be covered by technological measures through social measures while prioritizing decisions.

3.4 Introduction of models that address accountability

Consideration should be given to accountability, explainability, interpretability, transparency, etc., and the data used should be specified. Also, the introduction of interpretable models should be thought out.

– While it has become possible to build complex predictive models based on a large amounts of data, it has become too complex for developers to explain the outcomes of such models.

3.5 Other social values

Consideration should be given to social values other than fairness, accountability and transparency.

– It is important to give consideration to various social values.

– Consideration should be given to values such as democracy and well-being, and to make sure the design and content of services do not infringe on the diversity of values held by individuals.

4 Implementation phase

4.1 Principle of human involvement in the use of profiling and AI

Human involvement in the evaluation process using profiling and AI should be considered.

– Article 22 of the GDPR on fully automated decision-making sets forth the "human involvement principle" which states that, in principle, humans must be involved in decisions that have a significant impact on human lives.

– On the other hand, there is criticism that forcing human involvement will in fact lead to human alienation. If applying the principle of "human involvement" requires extensive human labour, it may encourage the use of proxy variables such as educational background and gender, thereby inducing discrimination.

– For the time being, in light of the risks inherent in fully automated decision-making, it is recommended that a reasonable discipline be explored through the development of appropriate procedures and systems; for example, by conducting a human review only when an appeal is filed against an automated decision based on profiling.

4.2 Accountability

When data utilization has an impact on individuals and stakeholders, consideration should be given to explaining the process behind that outcome in non-technical terms.

– There is still no public consensus on the appropriateness of the content and extent to which the explanation should be given.

– For the time being, it is necessary to provide justifications that evaluatees can understand after data processing, analysis and evaluation is completed (especially when responding to complaints).

4.3 Implementation of safety control measures

In light of the potential impact on individuals of data-processing results, careful attention should be paid to safety management.

– Implementation of safety management measures is required, taking into consideration the extent to which individuals would suffer from the infringement of rights and interests due to leakage, loss, or damage of data analysis results.

4.4 Accuracy of data content

The accuracy and authenticity of data content should be guaranteed.

– It is necessary to ensure the accuracy and authenticity of the input data used for data use, analysis and evaluation (including profiling).

– Methods include deleting data that is no longer needed and presenting annotations and data quality as meta-information, which is information about the reliability of the information.

4.5 Establishment and handling of procedures for disclosure, correction, suspension of use, withdrawal of consent, etc. of data analysis results and constructed health scores

Procedures for handling complaints from data subjects and evaluees should be developed.

- Input data used for data analysis and evaluation is subject to rights of disclosure, correction, suspension of use, etc., as long as that input data is categorized as “retained personal data”.
- As for the results of data analysis, the GDPR and the California Consumer Privacy Act (CCPA) treat profiling results (or “inferences” in the CCPA) as “personal information”. Companies should, to the extent possible, respond to requests for disclosure, correction, etc. in the same way for personal information as for input data.
- It is also necessary to ensure accessibility, taking into account the digital divide such as among elderly people.

4.6 Data handling in case of death

The handling of data in the event of the death of the data subject or after service suspension should be clearly stated when obtaining consent.

- While there is a legal basis that information about a deceased person is non-personal information, there has been a court case on whether bereaved family members should be allowed to access information on SNS, such as Facebook, that the deceased used before death, and the difficulty of making such judgements is being recognized.
- The will and reputation of the deceased should be respected and protected, and it is recommended that consideration be given before death such as by obtaining consent.

4.7 Improvement of literacy in data use, analysis and evaluation

Training should be provided to those who utilize data (such as business units and human resources departments of companies) to ensure they properly handle and evaluate the results of data processing and analysis.

- There is a concern about “automation bias”, in which humans overestimate and too easily accept decisions automated by AI and other technologies.
- Measures (such as training) should be taken to ensure that those who handle the results of data analysis do not blindly accept the results and “deify” the algorithms.
- In addition, consideration should be given to methods of communicating the results of data analysis so as not to hurt data subjects.

4.8 Responses of data subjects, service providers, etc.

Designers should consider potential influences on the behaviours and preferences of data subjects, those who use services, etc.

- Numbers that are overlooked in inputs (dark numbers) can lead to inappropriate outcomes.
- Such oversights can cause over-adaptation, gaming and loss of diversity.

4.9 Auditability

Auditability should be ensured.

- Trust in data users can be built if interested third parties understand in detail the processes of and risk measures against data use through information disclosure.
- New York City passed an ordinance on automated decision systems used by city agencies and established a task force to monitor the fairness of algorithms used by the city administration.

---

5 Data circulation phase

5.1 Provision of data to third parties and its governance with consideration to the impact on data subjects, etc.

Governance system should be agreed upon through contracts on data provision, etc., regarding the purpose of data use by those to whom the data is provided, restrictions on rolling distribution, prohibition of re-identification, etc.

- When providing data to a third party, it is recommended to not only confirm the legitimacy of the data provision but also consider the potential impact on data subjects of data utilization by third parties.

5.2 Procedures to provide the results of data analysis and constructed health scores to third parties

Consideration should be given when providing data analysis results to third parties.

- It is recommended that companies carefully consider whether to provide results of sensitive data analysis to third parties, such as credit scores.
- Whether the provision of sensitive information to third parties can be authorized through opt-out procedures should be thoroughly discussed in relation to reputation risk and legal basis.

5.3 Government access

Policies on how to respond to inquiries from the police and other national agencies should be established.

- It is worth considering establishing policies on how to respond to inquiries from the police and other national agencies, as well as from bar associations.

Keys to building public trust in the emerging era of data utilization

Data utilization in healthcare is still in its infancy. Abstract discussions on data use and privacy tend to be conservative and do not work constructively in emerging industries. Therefore, what is needed now to build public trust in data utilization to solve various social issues is communication and opinion-exchange among industry, academia, government and the public on issues and challenges that are likely to arise when each party carries out its role. Through such dialogue, gaps caused by various biases and differences in information literacy can be reduced. It is important that all relevant parties understand correctly what data users hope to achieve.

As an effort to build public trust, a governance framework that guarantees trust among a wide range of stakeholders has been proposed (white paper "Rebuilding Trust and Governance: Towards Data Free Flow with Trust (DFFT)").[11] In addition, there are increasing expectations for data circulation that ensures trustworthiness by using Fourth Industrial Revolution technologies for verification and trust-building (Trusted Web Council report).[12] We plan to further discuss these issues and present them again soon.

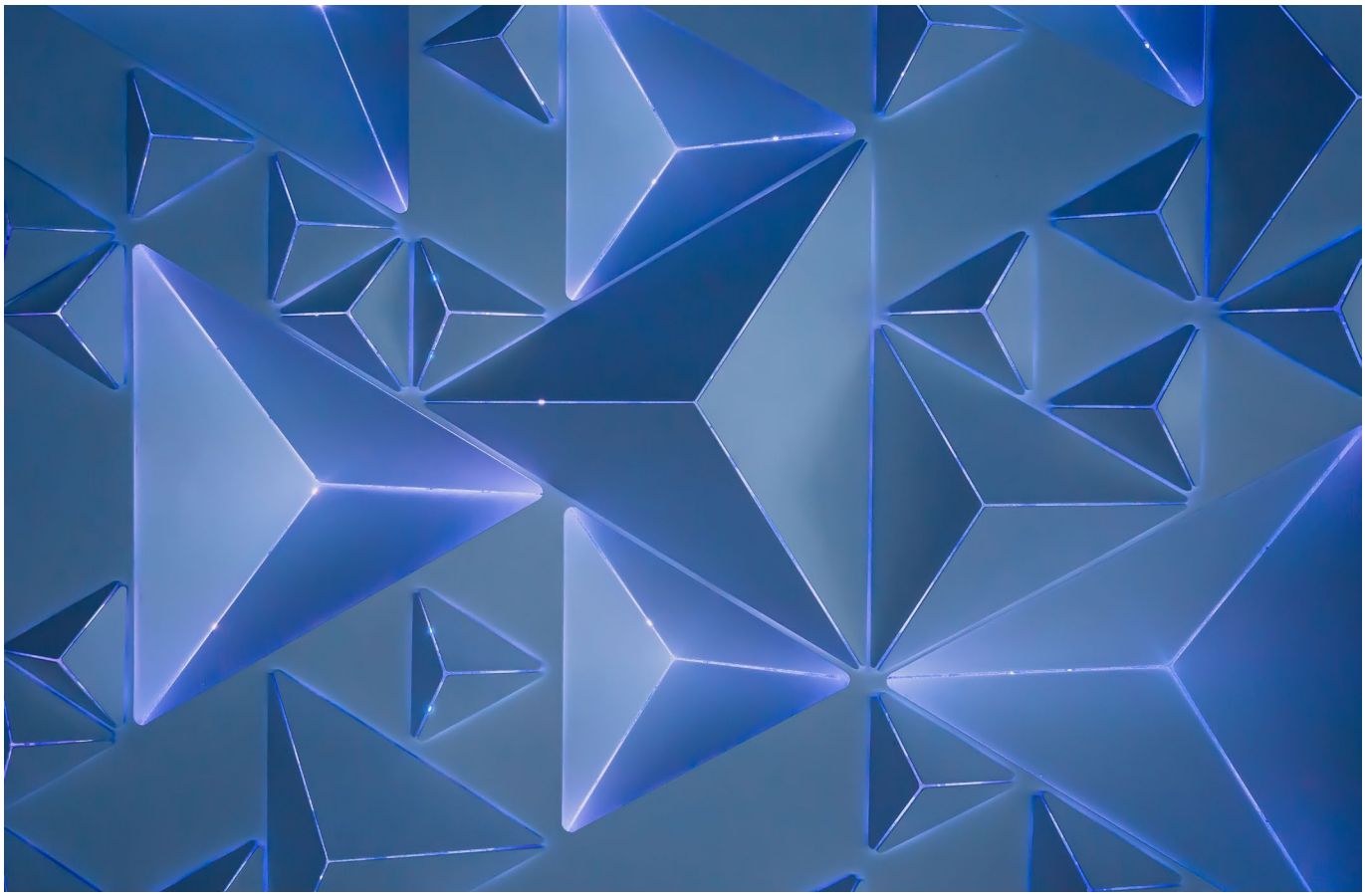
- **To reach a consensus among industry, academia, government and the public on the ideal vision of society we want to achieve through data utilization and the roles that each party should play**
- **To ensure that all stakeholders involved correctly understand the value that data users hope to achieve and the issues they want to solve**
- **To bridge gaps caused by various biases and differences in information literacy by exchanging opinions among industry, academia, government and the public on the issues and challenges likely to arise when each stakeholder carries out its role**

Conclusion

It is critical to understand that trust is something that is usually difficult to recognize and is often acutely felt only when it is lost. It is therefore a topic that is difficult to proactively address, especially in industries. This is something that humanity has already experienced in security. Security has been perceived as a cost centre by industries, and as laws and guidelines have been established in response to various problems, industry has taken a passive attitude: "Since it is a legal requirement, we will respond to it within the scope of the law." We should not repeat the same mistakes in data ethics. We should learn from the past.

As mentioned above, it will not be easy for the private sector to make collective efforts to gain public trust in data utilization in emerging technologies and industries. However, we believe these efforts will improve the sustainability of individual companies, the industries they belong to, and eventually the private sector as a whole. The most important thing is to communicate with the public by demonstrating through will, intention and attitude the value that data users want to create by using data. By doing so, companies can renew their business models and service designs and gain the public's trust.

The frameworks and issues presented in this paper are currently being further explored and discussed in depth. Conclusions will be published as a white paper in due course.



Contributors

Lead authors

Centre for the Fourth Industrial Revolution Japan,
World Economic Forum:

Seiichiro Yamamoto
Healthcare Data Policy, Project Lead

Takanori Fujita
Healthcare Data Policy, Project Lead

Yasunori Suzue
Healthcare Data Policy, Project Fellow/SOMPO Holdings

Fumiko Kudo
Project Strategy Lead

Jonathan Soble
Editorial and Communication Lead

Seiya Sasaki
Healthcare Data Policy, Intern

Sakiko Negishi
Healthcare Data Policy, Intern

Atsumi & Sakai:

Takafumi Ochiai
Partner Lawyer

Keio University School of Medicine:

Hiroaki Miyata
Professor

The World Economic Forum thanks the project community members for their insightful review and feedback.

Endnotes

1. "Society 5.0", *Cabinet Office*, https://www8.cao.go.jp/cstp/english/society5_0/index.html.

2. World Economic Forum, "The Great Reset", *The Great Reset Dialogues*, 2020, https://www.weforum.org/great-reset/

3. Organization for Economic Co-operation and Development, *The OECD Privacy Guidelines*, 2011, http://www.oecd.org/sti/ieconomy/49710223.pdf

4. Future of Life Institute, *Asilomar AI Principles*, 2017, https://futureoflife.org/ai-principles/

5. Organization for Economic Co-operation and Development, *Recommendation of the Council on Artificial Intelligence*, 2019, https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449

6. European Commission, *Ethics Guidelines for Trustworthy AI*, 2019, https://ec.europa.eu/futurium/en/ai-alliance-consultation/guidelines#Top

7. World Medical Association, *Declaration of Helsinki*, 2013, https://www.wma.net/what-we-do/medical-ethics/declaration-of-helsinki/

8. World Economic Forum, *Redesigning Data Privacy: Reimagining Notice & Consent for human technology interaction*, 2020, http://www3.weforum.org/docs/WEF_Redesigning_Data_Privacy_Report_2020.pdf

9. Solove, Daniel J., "Introduction: Privacy self-management and the consent dilemma." *Harvard Law Review*, Vol.126, 2012 pp.1880-1903

10. "Ethically Aligned Design in Practice", IEE Ethics in Action in Autonomous and Intelligent Systems, https://ethicsinaction.ieee.org/#series

11. World Economic Forum, Rebuilding Trust and Governance: Towards Data Free Flow with Trust (DFFT), 2021, https://www.weforum.org/whitepapers/rebuilding-trust-and-governance-towards-data-free-flow-with-trust-dfft

12. Trusted Web Promotion Council, Trusted Web White Paper (Draft), 2021, Documents/Trusted_Web_ホワイトペーパー（案）_v1.0.md at master · TrustedWebPromotionCouncil/Documents · GitHub