

Principles and Practice for the Ethics Use of AI in Hitachi's Social Innovation Business

The governance and ethics of AI are key issues for Hitachi, which recognizes the significant societal impact associated with the use of this technology across its extensive business domains. Many different aspects of AI ethics have already been highlighted, among them safety, privacy, fairness, transparency and accountability, and security. Meanwhile, not only the principles, but also their practical implementation are rapidly becoming an area of significant activity, examples of which include the publication of draft EU regulations on AI and METI's AI Governance Guidelines. Hitachi published its own "Principles guiding the ethical use of AI in Social Innovation Business" in February 2021. Hitachi is also using these principles as a basis for the practical implementation of initiatives to establish governance across all steps including research, PoC, product development, and maintenance. This article describes the ideas that underpin Hitachi's AI Ethical Principles and gives an overview of their practical implementation.

Jun Yoshida
Kohsuke Yanai, Ph.D.

1. Introduction

Artificial intelligence (AI) is becoming increasingly important as a source of innovation, being adopted in a wide range of fields including railways, energy, healthcare, manufacturing, and finance. Meanwhile, the challenges facing society are becoming severe and complex, including environmental problems, how to improve resilience, and how to enhance people's quality of life (QoL) at a time when their values are becoming ever more diverse. AI is expected to play a significant role in solving these societal challenges and will be used across widespread areas of society in the future.

With the launch of Lumada, Hitachi is accelerating the use of digital technologies in its Social Innovation Business. In Lumada, AI is an essential driving force behind the creation of new value. Hitachi's Social Innovation Business often deals with social infrastructure or services that have

a strong public interest component, and the impact of AI on society will be considerable. If society is to be both fair and trustworthy, it will require that AI be deployed, implemented, and operated appropriately based on an accurate understanding of its behavior and the associated risks.

Hitachi published its "Principles guiding the ethical use of AI in Social Innovation Business" (hereafter "AI Ethical Principles") in February 2021 as a corporate policy for development and social implementation of AI⁽¹⁾. This was also accompanied by the publication of a white paper describing Hitachi's initiatives in the area of AI ethics. These set out the standards of conduct and items to be addressed that are unique to Hitachi in the operation of its Social Innovation Business.

Hitachi has since its founding adopted a human-centric perspective. At Hitachi Mine, for example, the company's birthplace, a 156-m-high chimney was constructed for the smelter in 1915 to alleviate smoke pollution. At the time, it was the world's tallest chimney. Observation points were

also built in the nearby hills. The outcome of this was a management system that linked the state of the atmosphere with smelter production while also being open to the public. Hitachi was also among the first to incorporate ethics into its medical and neuroscientific research, becoming the first medical equipment manufacturer to establish an ethical review committee in 2000. Likewise with its consideration for privacy protection in the use of data, Hitachi in July 2014 appointed a personal data manager and set up a privacy protection advisory committee.

Hitachi's corporate mission is to "contribute to society through the development of superior, original technology and products," and as such ethics is deeply embedded in its culture, providing a steadfast underpinning to its development of technologies and products across all fields, not just AI. Hitachi's AI Ethical Principles are an expression of this in the context of AI and will serve to unify understanding of this topic across all staff and to guide the establishment of in-house governance systems.

The next section describes Hitachi's AI Ethical Principles and the ideas that underpin them. This is followed by a section explaining how they will be put into practice.

2. Hitachi's AI Ethical Principles

The Principles of Human-centric AI Society was published in Japan in March 2019. Internationally, the Organisation

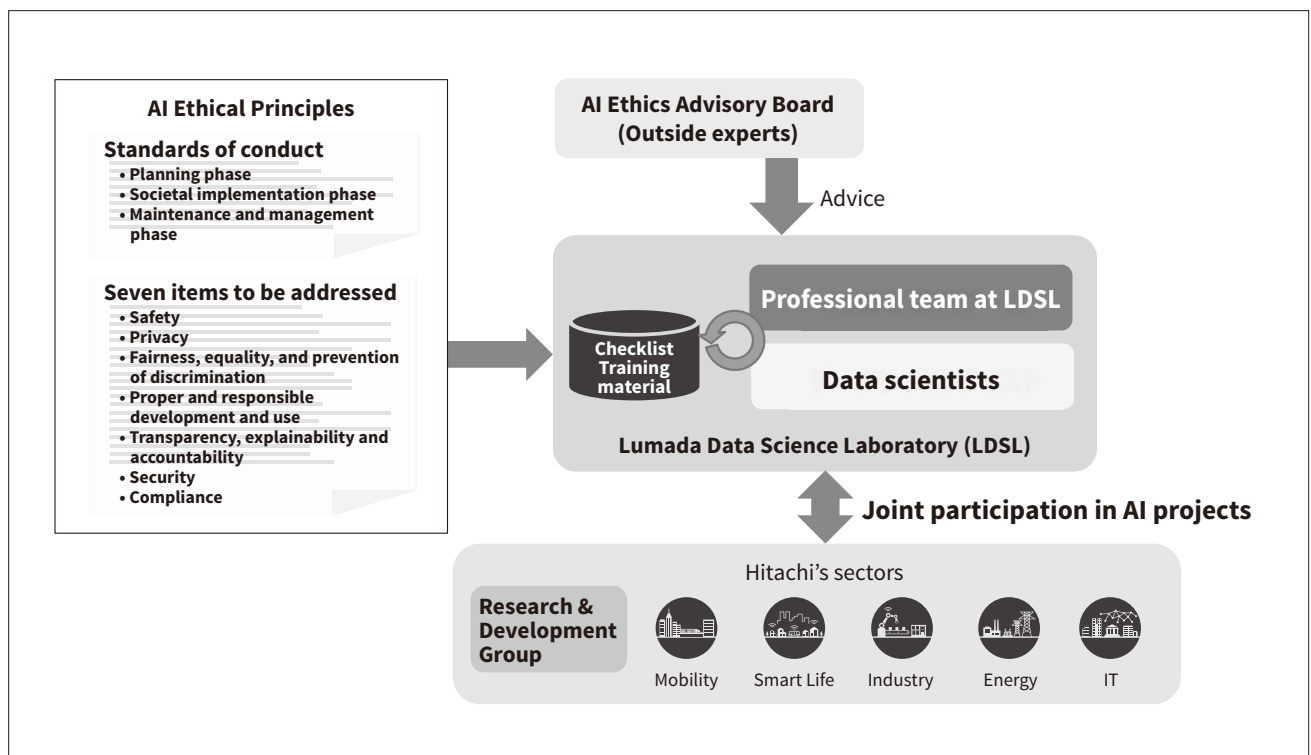
for Economic Co-operation and Development (OECD) adopted its Principles on Artificial Intelligence in May 2019 and the G20 adopted its own AI principles that draw on the OECD principles in June 2019. Various such national and corporate initiatives are also summarized in the Report 2021 of the Conference toward AI Network Society⁽²⁾. The problem of AI ethics has global recognition and the issues are being clarified through international debate that incorporates a diversity of viewpoints.

Hitachi, meanwhile, has formulated its own AI Ethical Principles that are informed by this international debate while also taking account of its corporate culture and the nature of its business that deals in social innovation. The principles identify standards of conduct in the three phases, the "planning phase," "societal implementation phase," and "maintenance and management phase," as well as seven "items to be addressed" that apply to all three phases (see Figure 1).

The process of formulating the AI Ethical Principles involved discussion across more than 10 business divisions to ensure that the outcome was in harmony with Hitachi's culture and operations. Trial risk assessments based on the principles were also undertaken by Hitachi research and development divisions in Japan, USA, Asia, and Europe to verify that they were workable in practice. Through this process, Hitachi was able to formulate principles that expressed concepts shared across the different business divisions and geographical regions and did so in a common language.

Figure 1 — Overview of Hitachi's Approach to AI Ethics

Hitachi's AI Ethical Principles stipulate standards of conduct across the "planning," "societal implementation," and "maintenance and management" phases of AI, and also seven "items to be addressed" that apply to all three phases.



Training material and a risk assessment checklist linked to the principles were also developed to facilitate the operation of AI governance (see **Figure 1**). The following sections describe the standards of conduct and items to be addressed stipulated in the AI Ethical Principles.

2. 1

Standards of Conduct in Each Phase

(1) Planning phase: Development and use of AI will be planned for the realization of a sustainable society

Hitachi's Social Innovation Business is forging new relationships in society and the economy while also being aware of the legal and other systems that govern how things are done⁽³⁾. Under this social challenges, it is not enough just to come up with an idea. Rather it involves working to develop concepts through a trial-and-error process of discovery and exploration in the public arena over a period of time. It is also essential that this time-extended conceptualization process includes planning for the use of AI to achieve appropriate goals from an ethical perspective.

Accordingly, Hitachi places an emphasis on addressing AI ethical considerations from the planning phase of its Social Innovation Business, not just at the stage of supplying services and solutions. AI ethics is put into practice from the earliest phases of innovation that encompass research and development and proof-of-concept (PoC).

(2) Social implementation phase: AI will be societally implemented with a human-centric perspective

Recognizing how large the impact on society will be from the use of AI in its Social Innovation Business, Hitachi understands the importance of implementing it in a responsible manner. This is a philosophy that Hitachi has followed since its founding.

Key features of AI that distinguish it from past technologies include that, (a) it is used for automation and the replacement of human workers, (b) its operation is determined not by explicit human instructions, but rather by the content of its training data, and (c) its operation is statistical by nature. The people involved need to have an in-depth understanding of these characteristics. Along with quality assurance for the AI algorithms and the systems that use AI, it is essential when applying the technology to use cases with a large societal impact that account also be taken of factors such as the validity of the training data and the transparency of the AI and associated systems. Adopting a human-centric perspective is fundamental to achieving this, and asking what the public actually needs should be the first question raised when putting AI into implementation.

(3) Maintenance and management phase: AI will be maintained and managed to provide long-term value

Social infrastructure remains in operation for decades. This obliges Hitachi to consider the implications of AI in its businesses over comparatively long periods, well beyond initial deployment.

For example, factors such as environmental change may result in changes in the data input to an AI over time, which would likely have implications for the system as a whole. This makes the maintenance and management of AI essential so as to ensure that systems incorporating the technology can continue to deliver value reliably over long time periods.

This involves more than just monitoring the AI for abnormal operation. It also means paying particular attention to the public benefits delivered by systems with AI and putting an effort into monitoring them to determine whether or not they continue to deliver what society needs.

2. 2

Seven Items to be Addressed

Hitachi's AI Ethical Principles identify seven "items to be addressed" that apply across the planning, societal implementation, and maintenance and management phases. These are: "Safety," "Privacy," "Fairness, equality, and prevention of discrimination," "Proper and responsible development and use," "Transparency, explainability and accountability," "Security," and "Compliance." They are also further broken down into specific checks. The phases of AI development include devising innovative new ideas, research and development, PoC, and solution and system development. The intention is that projects should consider these items at each of the different phases to verify that the company is acting in accordance with the AI Ethical Principles. The seven items can also be used as a basis for gathering and classifying high-risk cases that will be of use in areas like training or the formulation of risk mitigation measures. In this way, the seven items to be addressed will be central to Hitachi's efforts to put AI ethics into practice.

3. Practical Implementation

Hitachi has established a professional team of AI ethics specialists at the Lumada Data Science Lab. (LDSL), its center for AI and analytics. The team serves as a center of excellence (CoE) for AI ethics at Hitachi and includes data scientists and other staff with expertise in areas like privacy protection. It is responsible for establishing practices that facilitate compliance with the AI Ethical Principles as well as the continuous improvement of AI governance. The LDSL plays a central role in Hitachi's staff training for maintaining AI ethics and the research and development of relevant technologies. The LDSL is also a place where AI ethics are put into practice, being host to a large number of the leading co-creation projects undertaken by Hitachi's Social Innovation Business (see **Figure 1**).

Hitachi has also established an AI Ethics Advisory Board made up of outside experts to enable further improvements in its work on AI ethics based on objective assessment and

with reference to views from outside the company. The advisory board is able to offer advice not only on the technology of AI, but also from other perspectives such as that of consumers or social science.

3.1

Risk Management

Hitachi conducts risk assessments on individual projects using a checklist developed from the items to be addressed identified in the AI Ethical Principles, which include a focus on safety, fairness, transparency, explainability, and so on. This provides an efficient way to identify the risks that arise in AI research and development or in businesses that use AI. The checklist results provide a basis for advice on what to watch out for in project execution and also assist with finding countermeasures. Managers are also able to refer to the risks uncovered by the checklist when deciding whether or not work should proceed. The AI Ethics Advisory Board of outside experts, meanwhile, is available to discuss specific projects and provide advice as needed.

Conducting these risk assessments also facilitates the collection of information on wide-ranging AI research and development and on projects that utilize the technology, and this provides a basis for governance. Moreover, working through the risk assessment process allows for the collation and consolidation of knowledge about AI ethics.

The checklist-based risk assessments are conducted at various different phases in the AI lifecycle. In the research and development phase, for example, they can be used to screen research projects when at the beginning of the projects to verify the risks and whether it is appropriate to proceed. This checklist screening is conducted for more than 600 research projects each year. Checklist-based risk management is also used during commercial deployment and conducted for each of the phases from winning the customer order to development. Approximately 100 projects are reviewed in this way each year.

3.2

Quality Assurance of Services, Systems, and Products that Use AI

Growth in the development of services, systems, and products that use AI across Hitachi's businesses has prompted it to strengthen its AI-specific quality assurance practices. This has involved the formulation of development and quality assurance guidelines specifically for AI that are based on the Machine Learning Quality Management Guideline (AIQM)⁽⁴⁾ of the National Institute of Advanced Industrial Science and Technology and Hitachi's own AI Ethical Principles. The guidelines have been made available throughout Hitachi Group. The guidelines are a systematic compilation of topics that include the basic philosophies of AI quality assurance, how to develop AI models in the PoC and development phases, monitoring of prediction accuracy

during operation, how to handle retraining if accuracy deteriorates, the importance of safety in relation to AI ethics, privacy protection, and achieving fairness. Along with the guidelines being available to developers during the development of systems and products that use AI, they are also used to conduct reviews where they augment the wider quality assurance process by providing a particular focus on AI.

3.3

Training of AI Engineers and AI Ethics Education

An in-depth understanding of the AI technology is essential for complying with the AI Ethical Principles. Hitachi has established its own Data Science Community that brings together data scientists from across Hitachi Group to share their individual knowledge and know-how. The community is playing a central role in the training of approximately 3,000 data scientists with expertise in AI and data science. By training up this human capital to have an in-depth understanding of AI, Hitachi is establishing the foundations for its ethical use of AI.

Along with this AI training, Hitachi is also conducting practical training in AI ethics and taking the initiative to raise awareness of the subject within the company. This involves holding AI ethics study groups within the Data Science Community. Discussion groups are also held to keep people up to date with the latest developments in the field, covering topics such as what is happening with AI ethics in the public sphere as well as relevant research, internal governance, and regulation. These groups are open to all Hitachi staff regardless of profession or department and have been attended by more than 500 people, including senior management. The classroom training and study groups have also been augmented with workplace-based training that looks at specific use cases. Printed materials containing around 10 such AI use cases are distributed to workplaces where staff form groups of between 10 and 20 people, choosing several of use cases from this distributed resource material and then discussing the issues they raise in the context of AI ethics. More than 1,000 staff with an involvement in the research or use of AI have participated in this program.

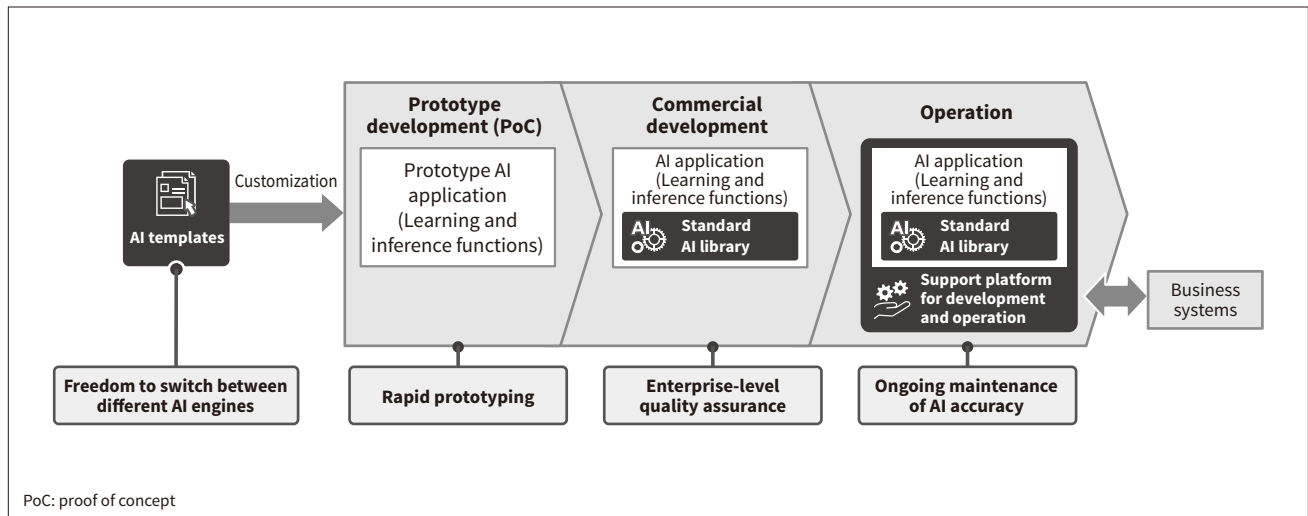
3.4

Research and Development of Technologies for Ensuring AI Ethics Compliance

In addition to in-house practices, Hitachi is also pursuing the research and development of technologies for ensuring AI ethics compliance based on its corporate mission of "contributing to society through the development of superior, original technology and products." Along with AI algorithms that make AI explainable or enhance fairness by detecting and minimizing data biases, the topics covered by this research and development also include ways of monitoring models and detecting performance degradation. Work is

Figure 2 — AI Application Framework

The framework consolidates Hitachi's expertise in the development and operation of AI and supports the development and operation of efficient, high-quality AI systems in highly specialized enterprise applications that demand a high level of knowledge.



also being done on data governance practices such as data lineage processes for achieving data traceability. Some of this research work is described in the article “Research and Development of AI Trust and Governance”⁽⁵⁾.

Hitachi also provides an AI application framework for faster AI implementation in corporate and social infrastructure applications that is based on the AI Ethical Principles⁽⁶⁾ (see **Figure 2**).

This framework consolidates Hitachi's expertise in the development and operation of AI and supports the development and operation of efficient, high-quality AI systems in highly specialized enterprise applications that demand a high level of knowledge. This provides ongoing monitoring and analysis of the input data and results once the AI system is in operation and the automatic detection of anomalous data or results based on predefined rules. This can prevent the degradation of prediction accuracy. If such degradation does occur, the AI application framework also provides the ability to call up a newly retrained model and to swap out the model used previously.

4. Conclusions

This article has described Hitachi's principles guiding the ethical use of AI in Social Innovation Business and how these principles are being put into practice.

By formulating its own set of principles based on its corporate culture and the nature of its business, Hitachi is seeking to document its stance on actions across the company and to achieve consistency in staff understanding of the subject. Work on establishing in-house governance systems has also accelerated. There is a clear sense that the task of formulating the principles has left staff with a greater appreciation than before of AI and its uses.

There is the view that AI governance and innovation are two different things and that strengthening AI governance will pose obstacles to innovation. However, governance and principles for using AI are essential if innovation is to happen and the research, development, and application of AI

Figure 3 — Kyōsō-no-mori Webinar

Hosted in October 2021 by the *Kyōsō-no-mori* facility in Kokubunji, Tokyo, the webinar was an opportunity to discuss AI governance with experts from both inside and outside of Hitachi.



can only proceed with the consent of the public based on international cooperation and agreement among stakeholders. Hitachi in the past has engaged in dialogue with stakeholders on such AI technologies as its complementary metal-oxide semiconductor (CMOS) Ising computer^{(7),(8)} and debating AI⁽⁹⁾. It is also striving to incorporate outside perspectives into its research and development, hosting the *Kyōsō-no-mori* Webinar⁽¹⁰⁾ on AI governance and trust in 2021 (see **Figure 3**). In the future, Hitachi intends both to enhance social, environmental, and economic value through a raised awareness of the importance of working in harmony with multiple stakeholders and to contribute to the creation of a comfortable and sustainable society based on human dignity and higher QoL around the world.

Authors



Jun Yoshida

Lumada Data Science Laboratory, Hitachi, Ltd.
Current work and research: Organizing a professional team at LDSL, customer co-creation using AI and analytics, and fostering data scientists within Hitachi.



Kohsuke Yanai, Ph.D.

Media Intelligent Processing Research Department, Center for Technology Innovation – Advanced Artificial Intelligence, Research & Development Group, Hitachi, Ltd. *Current work and research:* Research and development of media intelligence processing. *Society memberships:* The Japanese Society for Artificial Intelligence (JSAI).

References

- 1) Lumada Data Science Laboratory, “Ethical Use of AI” in Japanese, <https://www.hitachi.co.jp/products/it/lumada/about/ai/ldsl/index.html#p05>
- 2) Ministry of Internal Affairs and Communications, “The Conference toward AI Network Society 2021 Report: Promoting ‘the Safe, Secure, and Trustworthy Implementation of AI in Society’” in Japanese, https://www.soumu.go.jp/main_content/000761967.pdf
- 3) N. Konno and I. Nonaka, “Drawing a Big Picture: Methodology of Conceptual Imagination,” Nikkei Business Publications, Inc., Tokyo (Jul. 2018) in Japanese.
- 4) National Institute of Advanced Industrial Science and Technology, “Machine Learning Quality Management Guideline 2nd Japanese Edition” (Jul. 2021) in Japanese.
- 5) N. Uchida et al., “Research and Development of AI Trust and Governance,” *Hitachi Review*, 71, Special Issue, pp. 22–29 (Mar. 2021).
- 6) Hitachi, Ltd., “Hitachi AI Application Framework ‘Justware’” in Japanese, <https://www.hitachi.co.jp/products/it/appsvdiv/service/justware/ai-apfw/index.html>
- 7) M. Yamaoka et al., “CMOS Ising Computer to Help Optimize Social Infrastructure Systems,” *Hitachi Review*, 66, pp. 634–639 (Aug. 2017).
- 8) T. Nakamura et al., “Justware AI Application Framework for Development and Operation of AI Applications,” *Hitachi Review*, 71, pp.147–151 (Feb. 2022).
- 9) K. Yanai et al., “Debating Artificial Intelligence: Advanced Research into AI,” *Hitachi Review*, 65, pp. 151–155 (Jul. 2016).
- 10) Hitachi, Ltd., “*Kyōsō-no-Mori* Webinar,” <https://www.hitachi.com/rd/open/kyosonomori/collaboration/index.html>