

OECD AI Transparency Report

Organization: Hitachi, Ltd (JPN)

Reporting Period: Q4 2025

Published: November 13, 2025

Section 1 - Risk identification and evaluation

a. How does your organization define and/or classify different types of risks related to AI, such as unreasonable risks?

In the Principles guiding the ethical use of AI in Social Innovation Business (hereinafter, “Guiding Principles for the Ethical Use of AI”) of the Hitachi Group (hereinafter, “We”), the Standards of conduct are defined across three phases: Planning, Social Implementation, and Maintenance and Management. In addition, seven Items to be addressed across all phases are stipulated: Safety, Privacy, Fairness, Equality and Prevention of Discrimination, Proper and Responsible Development and Use, Transparency, Explainability and Accountability, Security, and Compliance. Risks, including unreasonable risks, are defined from the perspective of ensuring these items are properly addressed. Furthermore, the degree of risk is also assessed, and those that affect human life, fundamental human rights, or influence people’s emotions and thoughts are classified as high risk.

• Principles guiding the ethical use of AI in Social Innovation Business

https://www.hitachi.co.jp/products/it/lumada/about/ai/ldsl/document/ai_document_en.pdf

b. What practices does your organization use to identify and evaluate risks such as vulnerabilities, incidents, emerging risks and misuse, throughout the AI lifecycle?

We have established processes for risk management throughout the entire lifecycle—Planning, Social Implementation, and Maintenance and Management—as set forth in the Standards of conduct defined by the Guiding Principles for the Ethical Use of AI. These processes are integrated into business operations and include identifying risks including vulnerabilities, assessing the degree of risks, and mitigating them through appropriate countermeasures. Furthermore, with regard to incidents, processes are defined covering detection, information sharing, countermeasures, and prevention of recurrence.

c. Describe how your organization conducts testing (e.g., red-teaming) to evaluate the model's/system's fitness for moving beyond the development stage?

We do not develop or provide in-house LLMs, and therefore do not conduct third-party assessments of LLMs. In cases where we develop or provide AI-based systems, quality verification by a third-party organization (independent from the development process) and customer acceptance testing are conducted.

d. Does your organization use incident reports, including reports shared by other organizations, to help identify risks?

Yes

e. Are quantitative and/or qualitative risk evaluation metrics used and if yes, with what caveats? Does your organization make vulnerability and incident reporting mechanisms accessible to a diverse set of stakeholders? Does your organization have incentive programs for the responsible disclosure of risks, incidents and vulnerabilities?

We conduct both quantitative and qualitative assessments of risks associated with the use of AI, and implement countermeasures based on these assessments. In particular, quantitative evaluation metrics are established for aspects such as quality, accuracy, and bias.

A reporting process has been established for AI-related vulnerabilities and incidents to the AI Supervisory Committee, and critical cases are shared company-wide through the Information Infrastructure Division and the organization responsible for overseeing AI governance. In addition, reporting mechanisms are accessible to a diverse set of stakeholders through these established channels. While we do not have an incentive program for disclosing risks, incidents, and vulnerabilities, disclosures are made when legally required.

f. Is external independent expertise leveraged for the identification, assessment, and evaluation of risks and if yes, how? Does your organization have mechanisms to receive reports of risks, incidents or vulnerabilities by third parties?

We leverage the expertise of external experts, for example by inviting external advisors to the AI Supervisory Committee.

In addition, a contact point is publicly available on our website to receive reports from third parties regarding risks, incidents, or vulnerabilities.

<https://www.hitachi.co.jp/support/inquiry/index.html>

g. Does your organization contribute to the development of and/or use international technical standards or best practices for the identification, assessment, and evaluation of risks?

We have appointed three experts to ISO/IEC JTC 1/SC 42 Artificial Intelligence to contribute to the development of international standards. As one example, in ISO/IEC TR 5469 Functional safety and AI systems, we contributed to shaping the core concept related to functions that ensure the safety of AI-controlled equipment. As a best practice for implementing this concept, we have developed plant control technologies utilizing AI. Furthermore, with respect to fundamental standards related to AI such as ISO/IEC 22989 Artificial intelligence concepts and terminology and ISO/IEC 42001 AI management system, we have been involved not only in

international standardization but also in the development of national standards, thereby contributing to the promotion of international standards.

h. How does your organization collaborate with relevant stakeholders across sectors to assess and adopt risk mitigation measures to address risks, in particular systemic risks?

To appropriately address systemic risks, we, primarily led by the R&D division, conduct research on technical risks and monitor related trends. These insights are shared with the Quality Assurance Division and other relevant organizations to ensure appropriate risk mitigation when implementing AI in business operations. In addition, such information is also shared across industries through academic societies and industry associations.

Any further comments and for implementation documentation

No answer provided

Section 2 - Risk management and information security

a. What steps does your organization take to address risks and vulnerabilities across the AI lifecycle?

In alignment with the Guiding Principles for the Ethical Use of AI, we identify and address risks and vulnerabilities across the entire AI lifecycle. This is achieved by implementing measures such as: risk review and assessment by those responsible for AI-related projects in the Planning Phase, based on guidelines and checklists; implementation and quality assurance in the Social Implementation Phase in accordance with AI development guidelines; and, the organization of monitoring and management items for AI-related projects in the Maintenance and Management Phase, as well as the use of frameworks for efficient and high-quality operations.

b. How do testing measures inform actions to address identified risks?

We also conduct testing measures such as general risk assessments during the estimation stage of project development. In particular, when AI-related projects fall under high-risk classification, risk assessments are conducted through an “AI Ethics Risk Assessment” by a specialized committee.

c. When does testing take place in secure environments, if at all, and if it does, how?

Before releasing products or services incorporating AI systems, we conduct testing in a secure environment that is independent from the production environment.

d. How does your organization promote data quality and mitigate risks of harmful bias, including in training and data collection processes?

We ensure the quality of AI systems by conducting development that considers both the AI system development process and the assurance of data quality, while also mitigating risks of harmful bias. Specifically, in accordance with our internal quality assurance management standards, the appropriateness of the development process and the validity of training data are assessed using the “AI System Quality Assurance Process Checklist.”

e. How does your organization protect intellectual property, including copyright-protected content?

With respect to risks of copyright and other intellectual property infringements, we organize and publish internal guidelines and e-learning materials for all employees. These resources outline the risks associated with using third-party works and provide important precautions when utilizing AI-generated outputs. In particular, the guidelines stipulate that employees should not input data with the intention of generating output similar to third-party intellectual property, in order to prevent unintentional infringement of protected content. They also require employees to check in advance the terms of use of such works to confirm whether their use in input is prohibited or restricted. Furthermore, the guidelines recommend that output from generative AI should be accompanied by a clear indication that they were generated by AI.

f. How does your organization protect privacy? How does your organization guard against systems divulging confidential or sensitive data?

We protect privacy by establishing and operating a Personal Information Protection Management Systems (PMS) as along with an add-on Privacy Impact Assessment (PIA) framework. Under the PIA framework, we identify relevant operations based on criteria such as the “handling sensitive personal information,” and impose security measures throughout the entire data lifecycle. In certain cases—such as handling medical or financial information—stricter security measures are applied.

g. How does your organization implement AI-specific information security practices pertaining to operational and cyber/physical security?
i. How does your organization assess cybersecurity risks and implement policies to enhance the cybersecurity of advanced AI systems?ii. How does your organization protect against security risks the most valuable IP and trade secrets, for example by limiting access to proprietary and unreleased model weights? What measures are in place to ensure the storage of and work with model weights, algorithms, servers, datasets, or other relevant elements are managed in an appropriately secure environment, with limited access controls in place?iii. What is your organization's vulnerability management process? Does your organization take actions to address identified risks and vulnerabilities, including in collaboration with

other stakeholders?

- iv. How often are security measures reviewed?**
- v. Does your organization have an insider threat detection program?**

We identify AI-specific security risks organized by international standards and industry associations, and based on these risks, incorporate security-related items into the AI Ethics Checklist and the Guidelines for the Use of Generative AI. This ensures that appropriate security measures are implemented when using AI.

- The Information Security Division assesses cybersecurity risks in accordance with the Cybersecurity Management Guidelines of the Ministry of Economy, Trade and Industry (METI) and the NIST Cybersecurity Framework, and formulates organizational security policies and measures that also take into account emerging external threat trends. These policies are disseminated from headquarters to the security promotion divisions at each business site, which in turn implement security measures at their respective sites.
- We have established a Confidential Information Management System under which all information handled internally, including valuable IP and trade secrets, is classified by confidentiality level, and strict management is applied across the entire data lifecycle—including acquisition, storage, use, transfer, and disposal—according to the defined confidentiality levels. For systems, IT equipment, and cloud services (including AI) that handle such confidential information, security and data protection measures are assessed prior to use, and only data permitted by the level of protective measures in place is allowed to be handled, thereby ensuring data protection.
- We proactively collect information on external vulnerabilities. For high-urgency vulnerability information, company-wide instructions for countermeasures and inspections are issued and implemented. In addition, software configuration information of systems is registered and managed so that when software vulnerability information is published, affected systems can be identified and prompt information sharing and countermeasure instructions can be provided to the relevant system administrators, in collaboration with relevant stakeholders as needed.
- We conduct regular reviews of security measures at least once a year; however, when security incidents occur or highly urgent vulnerabilities are reported, we also carry out ad-hoc reviews of measures and implement emergency responses as necessary.
- To prevent insider misconduct, we have implemented filtering functions for the external transmission of digital data and access to external content, and system logs are recorded within internal IT systems to enable forensic response. In addition, we have established an internal whistleblowing system, and mutual monitoring among employees further contributes to the prevention of insider misconduct.

h. How does your organization address vulnerabilities, incidents, emerging risks?

We address the risks as follows:

Vulnerabilities: Prior to business application, risk assessments—including those related to vulnerabilities—are conducted in advance based on AI governance guidelines and checklists. When vulnerabilities are identified, each division implements countermeasures to mitigate the risks. Where necessary, the AI governance organization provides advice on risk identification and the development of countermeasures.

Incidents: In the event of AI-related incidents, responses are carried out in accordance with incident response documentation, in collaboration with the AI governance organization of each division.

Emerging Risks: With regard to emerging risks from new technologies such as AI agents, information is gathered primarily by the company-wide AI governance organization, and rules, guidelines, and related documents are revised as necessary.

Any further comments and for implementation documentation

No answer provided

Section 3 - Transparency reporting on advanced AI systems

a. Does your organization publish clear and understandable reports and/or technical documentation related to the capabilities, limitations, and domains of appropriate and inappropriate use of advanced AI systems?
i. How often are such reports usually updated?

On our website, the following information is publicly available:

- Principles guiding the ethical use of AI in Social Innovation Business

https://www.hitachi.co.jp/products/it/lumada/about/ai/ldsl/document/ai_document_en.pdf

- Hitachi's Initiatives for the Ethical Use of AI in the Social Innovation Business

https://www.hitachi.co.jp/products/it/lumada/about/ai/ldsl/document/ai_whitepaper_en.pdf

- Practical Initiatives for the "Principles guiding the ethical use of AI" in Hitachi's Research & Development Group

https://rd.hitachi.co.jp/_ct/17728022#c17728022_h7

- Hitachi's Principles guiding the ethical use of AI and Their Implementation for the Social Innovation Business

<https://www.hitachihyoron.com/jp/archive/2020s/2021/sp/sp01/index.html#toc-12>

- Justware AI Application Framework: Hitachi Enterprise Application Services

https://www.hitachi.co.jp/products/it/appsvdiv/service/justware/ai-apfw/index.html?_fsi=QQkQT4S8

- “An AI Society Develops as Diverse People Engage in Dialogue and Share Value” | 5th Co-Creation Forest Webinar “AI Governance” Program 3 “AI Trusted by Society”

https://linkingsociety.hitachi.co.jp/_ct/17498004

- New “Generative AI Center” to Promote the Utilization of Generative AI Inside and Outside the Company, Accelerating Value Creation in the Lumada Business and Enhancing Productivity

<https://www.hitachi.co.jp/New/cnews/month/2023/05/0515.html>

- Column: Trends in Generative AI and Hitachi’s Research Initiatives

<https://www.hitachihyoron.com/jp/archive/2020s/2024/01/16/index.html#toc-1>

- Hitachi Publishes the Book Practical Guide to Generative AI

https://digital-highlights.hitachi.co.jp/_ct/17686288

- Global Key Initiatives: Utilization of Generative AI

<https://www.hitachi.com/en/about/it/contents/technology-strategy/contents1/>

i. These reports will be updated as necessary.

ii. The new significant releases will be reflected as necessary.

iii. The published documents include the following information:

- Hitachi conducts assessments through its human rights due diligence processes.

https://www.hitachi.co.jp/sustainability/report/social/human_rights.html

- Hitachi continuously collects information related to privacy protection and discloses documents such as the following:

https://www.hitachi.co.jp/sustainability/report/governance/info_security.html#anc-06

b. How does your organization share information with a diverse set of stakeholders (other organizations, governments, civil society and academia, etc.) regarding the outcome of evaluations of risks and impacts related to an advanced AI system?

To improve our AI ethics initiatives by incorporating external perspectives, we share information on the results of risk and impact assessments related to advanced AI systems through collaboration with an AI Advisory Board composed of external experts, as well as other specialists.

Additionally, we participate in ISO/IEC JTC1/SC42 to contribute to the standardization of AI-related knowledge as well as engage in the development of legal frameworks and guidelines on AI with governments and other institutions.

c. Does your organization disclose privacy policies addressing the use of personal data, user prompts, and/or the outputs of advanced AI systems?

We disclose our privacy policy as follows.

<https://www.hitachi.co.jp/utility/privacy/index.html>

Guiding Principles for the Ethical Use of AI establish the appropriate handling of personal information and the protection of rights, including privacy, with respect to input data used for training, evaluation, and operation of AI systems, as well as the data output by AI.

d. Does your organization provide information about the sources of data used for the training of advanced AI systems, as appropriate, including information related to the sourcing of data annotation and enrichment?

We do not possess in-house proprietary large language model (LLM).

We primarily provide products and services to enterprises and government agencies, and we do not disclose data used for customer-specific tuning to any parties other than the customer. If requested by the customer, we provide appropriate information in accordance with their request.

e. Does your organization demonstrate transparency related to advanced AI systems through any other methods?

We promote research and development of explainable AI (XAI) to ensure the transparency of AI systems.

https://rd.hitachi.co.jp/_tags/%E8%AA%AC%E6%98%8E%E5%8F%AF%E8%83%BD%E3%81%AA/

For generative AI agents, we are developing technologies that support human verification by presenting the specific sources referenced for each element of the response.

<https://www.hitachihyoron.com/jp/archive/2020s/2025/01/15/index.html#toc-11>

Any further comments and for implementation documentation

No answer provided

Section 4 - Organizational governance, incident management and transparency

a. How has AI risk management been embedded in your organization governance framework? When and under what circumstances are policies updated?

We position AI risk as a cross-organizational issue that must be addressed. An AI governance organization operating under the Chief Legal Officer (CLO) establishes rules and guidelines. For individual cases, each department's AI governance team collaborates with on-site personnel to assess risks. Policies are reviewed as needed, considering changes in AI technology and risks, as well as developments in domestic and international policies and regulations.

b. Are relevant staff trained on your organization's governance policies and risk management practices? If so, how?

We provide e-learning to all employees covering AI-related risks, basic countermeasures, and key considerations. Additionally, we offer interactive training opportunities to ensure the proper implementation of AI risk management, including practical case studies. We also hold discussion

sessions to share and debate the latest global trends in AI ethics, research, internal governance, and regulations.

c. Does your organization communicate its risk management policies and practices with users and/or the public? If so, how?

We publicly disclose our AI risk management policies and practices on our corporate website, including:

- Principles Guiding the Ethical Use of AI for Social Innovation Business

https://www.hitachi.co.jp/products/it/lumada/about/ai/ldsl/document/ai_document_en.pdf

- Hitachi's Commitment to AI Ethics for Social Innovation Business

https://www.hitachi.co.jp/products/it/lumada/about/ai/ldsl/document/ai_whitepaper_en.pdf

- Practical Efforts Toward "Principles Guiding the Ethical Use of AI" by Hitachi R&D Group

https://rd.hitachi.co.jp/_ct/17728022#c17728022_h7

- Hitachi's Principles Guiding the Ethical Use of AI and Practices for Social Innovation Business

<https://www.hitachihyoron.com/jp/archive/2020s/2021/sp/sp01/index.html#toc-12>

- Justware AI Application Framework: Hitachi Enterprise Application Services

https://www.hitachi.co.jp/products/it/appsvdiv/service/justware/ai-apfw/index.html?_fsi=QQkQT4S8

- "A society with AI evolves through dialogue and shared values among diverse people" | Forest of Cooperative Creation Webinar #5 ""Governance of AI"" Program 3 ""AI trusted by society" Webinar: "AI Governance" – Program 3: "Trustworthy AI in Society"

https://linkingsociety.hitachi.co.jp/_ct/17498004

d. Are steps taken to address reported incidents documented and maintained internally? If so, how?

We document and maintain response rules for each type of incident. The company-wide AI governance organization reviews and updates these rules as needed, considering changes in AI technologies and risks, as well as developments in domestic and international policies and regulations.

e. How does your organization share relevant information about vulnerabilities, incidents, emerging risks, and misuse with others?

We promote awareness of AI governance by sharing information on regional AI regulations and service-specific risks via an internal employee portal. Regarding AI risks and vulnerabilities, we also share information externally through academic research and security-related organizations.

f. Does your organization share information, as appropriate, with relevant other stakeholders regarding advanced AI system incidents? If so, how? Does your organization share and report incident-related information publicly?

In the event of a major AI-related incident, our internal AI governance organization promptly shares information with relevant departments and, if necessary, coordinates with group companies and overseas offices. External reporting is conducted in accordance with laws and contracts. Such information is not publicly disclosed.

g. How does your organization share research and best practices on addressing or managing risk?

We share research and best practices related to AI risk response and management through the following methods:

Internally: Guidelines for AI usage outline risks and countermeasures by phase. Anticipated risks and countermeasures for specific cases are shared via an internal employee website.

Externally: Ongoing efforts toward Guiding Principles for the Ethical Use of AI are publicly disclosed on our external website.

https://rd.hitachi.co.jp/_ct/17728022

h. Does your organization use international technical standards or best practices for AI risk management and governance policies?

Our AI risk management and governance policies are developed with reference to international standards and best practices, including:

AIST's "Machine Learning Quality Management Guidelines"

METI and MIC's "AI Business Guidelines"

NIST's "AI Risk Management Framework"

ISO/IEC 42001

Any further comments and for implementation documentation

No answer provided

Section 5 - Content authentication & provenance mechanisms

a. What mechanisms, if any, does your organization put in place to allow users, where possible and appropriate, to know when they are interacting with an advanced AI system developed by your organization?

We primarily provide products and services to enterprises and government agencies. The usage and limitations of AI are explained in contracts and specification documents.

b. Does your organization use content provenance detection, labeling or watermarking mechanisms that enable users to identify content generated by advanced AI systems? If yes, how? Does your organization use international technical standards or best practices when developing or implementing content provenance?

We primarily provide products and services to enterprises and government agencies. In each contract and specification document, we explain the use of AI, the data utilized, and any applicable limitations.

We are also engaged in research and development aimed at enhancing content provenance, including the application of multi-layered digital watermarking to text generated by generative AI.

https://www.jstage.jst.go.jp/article/pjsai/JSAI2024/0/JSAI2024_4Xin281/_article/-char/ja

Any further comments and for implementation documentation

No answer provided

Section 6 - Research & investment to advance AI safety & mitigate societal risks

a. How does your organization advance research and investment related to the following: security, safety, bias and disinformation, fairness, explainability and interpretability, transparency, robustness, and/or trustworthiness of advanced AI systems?

We continuously promote research and development based on our Guiding Principles for the Ethical Use of AI. Examples include:

- AI governance for trustworthy AI

<https://www.hitachihyoron.com/jp/archive/2020s/2021/sp/sp01/index.html>

- Research supporting AI trust and governance

<https://www.hitachihyoron.com/jp/archive/2020s/2021/sp/sp02/index.html>

- "A society with AI evolves through dialogue and shared values among diverse people" | Forest of Cooperative Creation Webinar #5 "Governance of AI" Program 3 "AI trusted by society"

https://linkingsociety.hitachi.co.jp/_ct/17498004

- Practical Initiatives for "Principles Guiding the Ethical Use of AI" by Hitachi, Ltd. and the R&D Group

https://rd.hitachi.co.jp/_ct/17728022#c17728022_h7

b. How does your organization collaborate on and invest in research to advance the state of content authentication and provenance?

We conduct research and development to enhance content provenance, including applying multi-layered digital watermarking to generative AI output.

https://www.jstage.jst.go.jp/article/pjsai/JSAI2024/0/JSAI2024_4Xin281/_article/-char/ja

c. Does your organization participate in projects, collaborations, and investments in research that support the advancement of AI safety, security, and trustworthiness, as

well as risk evaluation and mitigation tools?

Our company is involved in the following research projects, collaborations, and investments.

- Statement to NIST's "U.S. Leadership in AI" plan
- Participation in ISO/IEC 42001 standardization
- Proposal activities for ISO/IEC TS 22440-1 (Functional safety and AI systems)
- Participation in Consortium of Quality Assurance for Artificial-Intelligence-based products and services (QA4AI Consortium) initiatives
- Member of MIC's "AI Network Society Promotion Council"
- Member of Cabinet Secretariat's "Digital Administrative Reform Council" and "Data Utilization System Review Committee"

d. What research or investment is your organization pursuing to minimize socio-economic and/or environmental risks from AI?

We established the "Hitachi-Tokyo University Lab" within the University of Tokyo as a central hub for innovation, under a collaborative industry-academia framework. This initiative aims to create a vision for realizing a "Super Smart Society" (Society 5.0) that brings prosperity to humanity.

The lab consists of two main projects:

"Habitat Innovation," which focuses on developing smart city concepts suitable for Society 5.0 and implementing them in real urban environments.

"Energy," which explores energy systems that support Society 5.0 from multiple perspectives. These projects are organically linked to promote joint research. The outcomes and findings of these collaborations are actively shared with society through open forums and other communication channels.

<https://www.ht-lab.ducr.u-tokyo.ac.jp/en/>

Any further comments and for implementation documentation

No answer provided

Section 7 - Advancing human and global interests

a. What research or investment is your organization pursuing to maximize socio-economic and environmental benefits from AI? Please provide examples.

We aim to realize social innovation through research and development of advanced technologies utilizing AI. We are deepening technologies such as generative AI, machine learning, knowledge processing and reasoning, explainable AI, multimodal AI, and XR/metaverse, targeting diverse data types including textual, audio, visual, and sensor data. We also promote research that considers AI transparency and fairness, aiming to build a trusted technological foundation. By enabling AI-driven prediction, decision-making, optimization, and support for judgment, we seamlessly connect technology development to business creation across sectors such as

Mobility, Energy, Connected Industries, and Digital Systems & Services—realizing a new society where AI and humans co-evolve.

<https://www.hitachi.co.jp/rd/careers/lab/ai/>

b. Does your organization support any digital literacy, education or training initiatives to improve user awareness and/or help people understand the nature, capabilities, limitations and impacts of advanced AI systems? Please provide examples.

We support digital literacy and education through:

- AI ethics training for all employees, Internal awareness activities, Discussion sessions on global AI ethics trends, research, governance, and regulations

https://rd.hitachi.co.jp/_ct/17728022

c. Does your organization prioritize AI projects for responsible stewardship of trustworthy and human-centric AI in support of the UN Sustainable Development Goals? Please provide examples.

We are advancing the use of AI to enhance societal sustainability through initiatives such as the AI projects outlined below.

- Correlation analysis and predictive diagnostics using disaster data and Hitachi Insights
- Smart charging solutions using AI

https://www.hitachi.co.jp/sustainability/download/pdf/ja_sustainability2024.pdf

d. Does your organization collaborate with civil society and community groups to identify and develop AI solutions in support of the UN Sustainable Development Goals and to address the world's greatest challenges? Please provide examples.

- We developed a “Policy Recommendation AI” that illustrates branching structures and factors among various possible future scenarios based on social models. In collaboration with experts from Kyoto University, we are analyzing Japan’s sustainability in the year 2050.

<https://www.hitachihyoron.com/jp/archive/2010s/2019/03/05c04/index.html>

- We are working to enhance the sustainability of social infrastructure by utilizing AI in its maintenance and operation.

https://www.hitachi.co.jp/products/it/society/catalog/whitepaper_AforSociety.pdf

- Hitachi’s metaverse, AI, and robotics initiatives support frontline workers.

<https://www.hitachihyoron.com/jp/archive/2020s/2025/01/15/index.html#toc-1>

Any further comments and for implementation documentation

No answer provided

