

# OECD AI Transparency Report

Organization: Hitachi, Ltd. (JPN)      Reporting Period: Q2 2026      Published: October 2, 2026

## Section 0 – Your role in the AI system lifecycle

Responding as (please select all that apply):

- Deployer
- Application developer (/provider)

If any of the fields above are selected, please indicate your organization size:

Large enterprise: 250 or more employees

## Section 1 – Risk Identification and Evaluation

1. How does your organization define and/or classify different types of risks related to AI?<sup>1</sup>

- Maintain an internal taxonomy, classification system, and/or set of risk area definitions
- Define enhanced safety measures for specific dangerous capability thresholds, that, once reached, require safeguards (e.g. capability threshold for dangerous Chemical, Biological, Radiological and Nuclear (CBRN), cybersecurity attacks, dangerous advanced autonomy capabilities, or other capabilities).
- Consult with external stakeholders (e.g., from academia, industry groups, civil society, or standard development organisations) when developing or updating risk taxonomies

How does your organization define thresholds at which severe risks, unless adequately mitigated, may be considered "unreasonable"? Please describe the framework(s) most relevant to your organization's role and context. If your organisation does not use formal thresholds, you may describe instead the policies, decision criteria, or governance processes you apply when assessing severe AI risks. You may also describe: • Specific examples of contexts in which your models or systems could, in your assessment, pose unreasonable or intolerable risk; • How mitigations are graduated by capability or propensity level; • How your organisation considers the specificities of different AI technologies, such as generative AI and agentic AI, including any specific risk assessment or governance mechanisms applied to agentic AI capabilities.

The Hitachi Group (hereinafter, “we”) has established criteria for high-risk AI, taking into account impacts on human rights, human life, safety, and other matters, with reference to the EU AI Act and other sources. As we consider AI risks from multiple perspectives, we do not rely exclusively on any single framework.

We adopt a risk-based approach based on the Guiding Principles for the Ethical Use of AI. In addition to the Standards of conduct for each of the “Planning,” “Social Implementation,” and “Maintenance and

Management” phases, the Guiding Principles set forth the following seven Items to be addressed: “Safety,” “Privacy,” “Fairness, Equality, and Prevention of discrimination,” “Proper and responsible development and use,” “Transparency, Explainability and Accountability,” “Security,” and “Compliance.” From these perspectives, we assess the purpose and context of AI use and treat as high risk any AI use that may have a significant impact on human life, fundamental human rights, individual dignity, or individuals’ autonomous decision-making. Where a risk is determined to be high, we consider and implement necessary risk mitigation measures and determine whether the AI may be used, taking into account the results of those measures.

**2. Both before and after deployment, what practices does your organization use to identify and evaluate vulnerabilities, incidents, emerging risks and misuse throughout the AI system lifecycle?**

- Adversarial testing to detect technical vulnerabilities
- Internal model, system, or application-independent research initiatives
- Vulnerability reports
- Incident reports, including reports shared by other organizations
- Publicly available benchmarks and measurement tools
- Internal benchmarks, standard metrics (e.g., frequency, severity scores)
- Structured techniques to explore the behavior or performance of a model and/or application, such as model elicitation or bottleneck assessments
- Qualitative assessments (e.g., expert judgment, narrative scoring)
- Impact assessments such as personal data protection assessment and fundamental rights impact assessments
- Specific risk identification practices for vulnerable groups

**Please briefly describe your organization's decision-making process for deploying an AI model or system including: who has decision authority; what evidence and criteria are required (e.g., performance, safety/security, robustness, fairness, privacy, compliance, independent external testing, evaluation, validation and verification (TEVV) or assurance); contextual caveats or limitations required in interpreting risk evaluation metric results; what approvals/sign-offs are needed; and what safeguards (e.g., mitigations, monitoring, or phased release) must be in place to move into deployment.**

We conduct risk management during each of the “Planning,” “Social Implementation,” and “Maintenance and Management” phases, based on the Guiding Principles for the Ethical Use of AI. For projects involving the use of AI, we assess risks from the perspectives of performance, safety, security, privacy, fairness, compliance with laws and regulations, and other relevant considerations. Where necessary, the relevant AI Governance Committee or an equivalent function conducts a review. Identified risks are mitigated through technical and institutional measures, and, taking the results of such mitigation into account, the person responsible for the relevant AI-related business activity determines whether the AI may be introduced. We also conduct continuous monitoring after introduction and implement necessary corrective actions if new risks are identified.

**Following deployment, how do you use incident reports — including those shared by other organizations — to identify emerging issues, update risk assessments, and trigger corrective actions.**

We have established processes for reporting information on AI-related vulnerabilities and incidents, including cases shared by other organizations, to the AI Governance Supervisory Committee. We analyze the reported cases to identify any new risks and assess their impact on existing risk assessments. We also share significant cases across the Hitachi Group through our internal information infrastructure functions and the AI Governance Supervisory Committee, and, as necessary, take corrective actions such as issuing alerts and reviewing guidelines and operational practices.

### 3. Does your organization have mechanisms to receive and/or share reports of risks, incidents or vulnerabilities by/with third parties?

Yes, both receiving and sharing

If yes, please indicate which stakeholders reporting mechanisms may be accessible to, recognising specifics will depend on the nature of the report and risk level. For each, indicate whether access is on a standing basis (Always), activated based on context (Ad hoc/as relevant), or not currently accessible.

|                                                                                                                                                                                                                                                                                                                                                                   | Always                   | Ad hoc /<br>as<br>relevant | Not<br>currently<br>accessible |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------|----------------------------|--------------------------------|
| Relevant employees (security, legal/compliance, etc.)                                                                                                                                                                                                                                                                                                             | <input type="checkbox"/> |                            |                                |
| Public authorities (e.g., public reporting offices for security incidents or data breaches)                                                                                                                                                                                                                                                                       |                          | <input type="checkbox"/>   |                                |
| Supervisory authorities (e.g., data protection authorities)                                                                                                                                                                                                                                                                                                       |                          | <input type="checkbox"/>   |                                |
| Industry partners or industry fora                                                                                                                                                                                                                                                                                                                                |                          | <input type="checkbox"/>   |                                |
| Users (end users, enterprise customers, developers using APIs)                                                                                                                                                                                                                                                                                                    | <input type="checkbox"/> |                            |                                |
| Business customers / deploying organizations (if different from "users")                                                                                                                                                                                                                                                                                          | <input type="checkbox"/> |                            |                                |
| Supply-chain partners (cloud, data providers, etc.): <i>The Hitachi Incident Response Team continuously accepts reports from third parties concerning risks, incidents, and vulnerabilities through its external contact point (<a href="https://www8.hitachi.co.jp/inquiry/it/hirt/jp/form.jsp">https://www8.hitachi.co.jp/inquiry/it/hirt/jp/form.jsp</a>).</i> | <input type="checkbox"/> |                            |                                |
| Academia / researchers                                                                                                                                                                                                                                                                                                                                            |                          | <input type="checkbox"/>   |                                |
| Independent auditors                                                                                                                                                                                                                                                                                                                                              |                          | <input type="checkbox"/>   |                                |
| Standards bodies / multi-stakeholder initiatives (OECD.AI, ISO/IEC, IEEE, etc.)                                                                                                                                                                                                                                                                                   |                          | <input type="checkbox"/>   |                                |
| Specific civil society organizations                                                                                                                                                                                                                                                                                                                              |                          | <input type="checkbox"/>   |                                |
| Media / general public (people or communities indirectly impacted)                                                                                                                                                                                                                                                                                                |                          | <input type="checkbox"/>   |                                |
| Other                                                                                                                                                                                                                                                                                                                                                             |                          |                            |                                |

**Please describe how your organisation determines when and how reporting mechanisms are made accessible to different stakeholders (e.g. based on legal or regulatory requirements, internal policies, or risk level). Please also provide public URLs or details on how to submit reports. If none, write “Not applicable”**

Public authorities and Supervisory authorities are notified where reporting obligations arise under applicable laws and regulations.

Industry partners or industry fora are notified where reporting obligations arise under industry practices or forum rules.

Reporting channels are kept open at all times for Relevant employees in accordance with internal rules and the risk management structure.

Reports from third parties concerning risks, incidents, and vulnerabilities are accepted at all times from Users, Business customers / deploying organizations, and Supply-chain partners

(<https://www8.hitachi.co.jp/inquiry/it/hirt/jp/form.jsp>).

We communicate, as appropriate, with Academia / researchers, Independent auditors, Standards bodies / multi-stakeholder initiatives, and Specific civil society organizations when the AI Governance Supervisory Committee requires external expertise.

For the Media / general public, we consider appropriate disclosure of information where a serious incident occurs or a societal impact is anticipated.

**How does your organization share relevant information as appropriate?**

- Publicly
- Information-sharing partnerships (with confidentiality agreements as appropriate)
- Reporting to regulatory authorities as relevant

**Please describe what type of information is shared.**

As part of its public information sharing, Hitachi has published the Guiding Principles for the Ethical Use of AI and communicates to society its initiatives based on the seven items to be addressed, including safety and privacy.

As part of its information-sharing partnerships, Hitachi invites external experts to the AI Governance Supervisory Committee to share their expertise. Hitachi also participates in CSIRT industry forums, where information on risks, incidents, and vulnerability reports, including those related to AI, is shared. Furthermore, when collaborating with customers and partner companies, Hitachi appropriately shares information after entering into confidentiality agreements where necessary.

With regard to reporting to regulatory authorities, Hitachi identifies compliance with laws and regulations as one of the items to be addressed and submits appropriate reports in accordance with applicable laws and regulations.

**4. If your organization has programs to incentivize risk(s) disclosure, incidents and systemic vulnerabilities, please share what type of programs you use?**

- Vulnerability Disclosure Programs (VDPs) with safe harbor / liability protection
- Recognition and reputation incentives
- Internal incentives and culture

Please provide public URLs and/or details on your incentive program(s). If none, write “Not applicable.”

The page describing the vulnerability disclosure process for the VDP is available here.

<https://www.hitachi.com/ja-jp/hirt/publications/hirt-pub/hirt-pub10008/>

The “Acknowledgments” page for recognition and reputation incentives is available here.

<https://www.hitachi.com/en/hirt/security/acknowledgments/>

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

Contributes

If contributes or uses, please select all that apply:

- ISO/IEC standards (e.g., ISO/IEC 27001, ISO/IEC 42001, ISO/IEC 23894, ISO/IEC 42005)

**6. Does your organization collaborate with relevant stakeholders across sectors to assess and adopt risk mitigation measures to address risks that could have broad and potentially cascading socio-economic impacts? Please indicate which stakeholders you collaborate with, and whether collaboration is on a standing basis (Always-on), activated based on context (Ad hoc/as relevant), or there is no current collaboration.**

|                                                                                                                                    | Always-on                | Ad hoc / as relevant     | No current collaboration |
|------------------------------------------------------------------------------------------------------------------------------------|--------------------------|--------------------------|--------------------------|
| Relevant employees (security, legal/compliance, etc.)                                                                              | <input type="checkbox"/> |                          |                          |
| Public authorities / regulators / AI safety institutes / other public sector bodies and inter-governmental coordination structures |                          | <input type="checkbox"/> |                          |
| Industry partners or industry fora                                                                                                 |                          | <input type="checkbox"/> |                          |
| Users (end users, enterprise customers, developers using APIs)                                                                     |                          |                          |                          |
| Business customers / deploying organizations (if different from “users”)                                                           |                          |                          |                          |
| Supply-chain partners (cloud, data providers, etc.)                                                                                |                          |                          |                          |
| Academia / researchers                                                                                                             |                          | <input type="checkbox"/> |                          |
| Independent auditors                                                                                                               |                          |                          |                          |
| Standards bodies / multi-stakeholder initiatives (OECD.AI; ISO/IEC; IEEE; AIID etc.)                                               |                          | <input type="checkbox"/> |                          |
| Independent civil society organizations (e.g., for evaluations, audits, or policy review)                                          |                          |                          |                          |

|                                                                    | Always-on | Ad hoc / as relevant | No current collaboration |
|--------------------------------------------------------------------|-----------|----------------------|--------------------------|
| Media / general public (people or communities indirectly impacted) |           |                      |                          |
| Other(s) (please add any other relevant stakeholder groups):       |           |                      |                          |
| Not applicable                                                     |           |                      |                          |

**Please describe examples of how your organization collaborates. Examples may include: participating in multi-stakeholder or industry initiatives; partnering with academia, civil society, or the public sector; establishing internal or joint governance structures with external experts or in partnership with other efforts to share knowledge and jointly address risks.**

Representatives from organizations across the Hitachi Group participate in the AI Governance Supervisory Committee, enabling ongoing collaboration among these organizations.

We also invite external advisors to participate in the Committee, thereby incorporating the expertise of external experts into our AI governance activities.

Furthermore, we have dispatched three experts to ISO/IEC JTC 1/SC 42 Artificial Intelligence, contributing to the development of international technical standards and best practices for the identification, evaluation, and assessment of AI risks. For example, we participated in the development of ISO/IEC TR 5469 Functional safety and AI systems and contributed to establishing core concepts for the safe application of AI to control systems. We have also been involved in the international standardization and development of national standards for ISO/IEC 22989 Artificial intelligence concepts and terminology and ISO/IEC 42001 AI management system, thereby supporting the dissemination of international technical standards.

In addition, we have made an inquiry contact point available on our website through which reports of risks, incidents, or vulnerabilities from third parties can be received.

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

**Any further comments and implementation documentation.**

*No answer provided*

## Section 2 – Risk Management and Information Security

**7. Across the AI system lifecycle, what steps does your organization take to manage risks and vulnerabilities? Please indicate which specific practices, tools or metrics your organization uses at each lifecycle stage. You may refer to tools or metrics present in the OECD.AI Catalogue of Tools and Metrics or suggest additions. Respondents may complete only the lifecycle stages relevant to their organisation's role. Stages that do not apply may be left blank.**

Plan and design

**Collect and process data**

**Build, integrate and interpret model**

**Verify and validate**

**Deploy**

**Operate and monitor**

**Please provide a consolidated description of your organisation's approach to risk management across the AI system lifecycle, including how the practices identified above operate together and any cross-cutting governance, review or escalation processes. Please also indicate how often these practices are reviewed and what types of events trigger reviews and updates. Where relevant, please distinguish between different types of AI systems (e.g. generative or agentic AI).**

In alignment with the Guiding Principles for the Ethical Use of AI, we identify and address risks and vulnerabilities across the entire AI system lifecycle by implementing measures such as: risk review and assessment by those responsible for AI-related projects during the research planning phase and the service planning phase, based on guidelines, checklists, and other relevant materials; 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 operations in the Maintenance and Management Phase, as well as the use of frameworks to achieve efficient and high-quality operations.

**Please describe any measures used to ensure data quality and fairness across the data collection and curation stage (e.g., dataset audits, filtering training data to remove harmful content or personal information, fine-tuning on curated data, use of moderation tools, assessment of limitations to model generalisability).**

We ensure the quality of AI systems by conducting development that considers both the AI system development process and the assurance of data quality. 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." In our research and development activities, we have also developed an "AI Ethics Checklist" and conduct assessments from the perspectives of Safety, Privacy, Fairness, Equality and Prevention of Discrimination, Proper and Responsible Development and Use, Transparency, Explainability and Accountability, Security, and Compliance.

## **8. What security measures are used in testing AI systems?**

- Isolated and secure test environments
- Network isolation from external access
- Locally hosted testing environments
- Restrictions on outbound Internet or API access, including a list of permitted domains or APIs
- Controlled environments for red-teaming and adversarial testing
- Isolation of staging environments (from production / operational environments)
- Controlled security vulnerability testing

- Human-in-the-loop controls for higher-risk test cases
- Access controls, including roles-based permissions
- Audit logging and monitoring
- Use of “AI-safe” test-data, as appropriate
- Use of synthetic, anonymized, or dummy data, as appropriate

**8.A. Please indicate, if relevant, the team or role responsible for overseeing secure testing and whether this responsibility is shared across multiple teams and/or functions.**

We have established a structure in which multiple teams work together and share responsibility for overseeing the secure testing of AI systems.

When a business operation in which an AI system is applied is classified as high risk, the AI Governance Supervisory Committee conducts a risk assessment through an “AI Risk Assessment” and oversees testing.

For the development of AI systems for internal use, the development division tests whether the systems meet the IT security standards established by the IT and information security divisions.

For the development of products or services incorporating AI systems, the development team conducts risk assessments in accordance with the standards for secure system development and operation, and conducts testing in a secure environment that is independent from the production environment. The Quality Assurance Division verifies security as an aspect of quality and oversees customer acceptance testing.

Through this structure, responsibility for overseeing secure testing is distributed and shared across multiple stages and organizational functions to minimize risks.

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

**9.A. Safeguards that promote lawful data use and transparency in AI training:**

- Maintain a copyright policy that addresses data sourcing, content provenance, safeguards, and organizational roles and responsibilities
- Refrain from the unlawful circumvention of technological protection measures in data sourcing
- Conduct data cleaning, qualification, and/or tokenization activities as part of a transformative data preparation process
- Consult and/or retain an intellectual property attorney to protect, enforce, and comply with, intellectual property rights.

**9.B. Safeguards that reduce the risk of dissemination of outputs that may infringe copyright:**

- Restricting outputs that raise ethical, human rights, or legal concerns—such as non-consensual intimate imagery, unauthorized deepfakes, or outputs that are substantially similar to copyrighted works.

**10. How do you protect privacy and prevent your systems from disclosing confidential or sensitive data across the AI lifecycle?**

- Make information about data privacy, practices, and controls accessible to a diverse set of stakeholders

- Use security controls, including encryption at rest and in transit
- Provide organizational customers with controls to configure and implement privacy choices (e.g., data retention, logging, access, or processing settings)
- Apply Privacy-Enhancing Technologies (PETs) where applicable (e.g. differential privacy, federated learning, synthetic data generation, homomorphic encryption)

**Describe key privacy measures and data-protection standards applied.**

We have established and operate a Personal Information Protection Management System (PMS) and a Privacy Impact Assessment (PIA) framework. Under the PIA framework, we identify operations that handle sensitive personal information and implement security measures throughout the entire data lifecycle. When handling medical or financial information, stricter security measures are applied.

As a personal information protection standard, we have established a Personal Information Protection Management System based on JIS Q 15001. In establishing the PIA framework, we referred to ISO/IEC 29100 and other relevant standards.

We have also established a Confidential Information Management System under which all information handled internally is classified by confidentiality level, and strict management is applied at each stage of acquisition, storage, use, transfer, and disposal according to the defined confidentiality levels. Before IT equipment and cloud services, including AI systems, are put into use, their security measures and data protection measures are assessed. Data is protected by allowing only data at confidentiality levels appropriate to the measures in place to be handled.

As a confidential information protection standard, our business sites obtain certifications such as ISO/IEC 27001.

**11. How does your organization assess cybersecurity risks and implement policies to enhance the cybersecurity of advanced AI systems across the AI system lifecycle? Respondents may complete the checkbox subsections in 11.A–11.D or describe their overall approach under 11.E.**

**11.A. Governance, policy and assessments**

- Conduct regular cybersecurity risk and vulnerability assessments tailored to AI models or systems
- Perform threat modeling that includes AI-specific risks (e.g., data poisoning, prompt injection)
- Maintain or update security policies aligned with recognized standards (e.g., NIST CSF, ISO/IEC 27001, SOC 2, CIS Critical Security Controls), and review them regularly in response to emerging risks or regulatory guidance
- Maintain vulnerability management processes (scanning, channels for external reporting, patching, remediation)
- Implement incident response, disaster recovery, or business continuity plans, including escalation procedures for high-severity incidents
- Train staff on cybersecurity best practices and AI-specific threats, including how to recognize and respond to misuse or unintended outcomes of advanced AI systems

**11.B. Technical and operational controls**

- Apply multi-factor authentication (MFA) and role-based access controls for advanced AI systems
- Encrypt AI assets (datasets, models, configurations) in storage and in transit as appropriate

- Restrict physical access to servers, GPUs, or devices hosting AI models, systems, and/or applications
- Monitor system activity and log access to detect anomalies or misuse

#### **11.C. Protection of sensitive or high value AI assets**

- Limit access to proprietary AI assets based on role or need-to-know
- Store model weights, datasets, and algorithms in secure, access-controlled environments
- Require NDAs or confidentiality agreements for staff and contractors with access

#### **11.D. Insider risk protection detections**

- Conduct pre-employment screening or role-sensitive vetting
- Provide staff training on cybersecurity best practices and AI-specific threats
- Implement insider-risk escalation or response procedures

#### **11.E. Please describe the main security measures your organization applies and indicate which are specific to particularly sensitive AI assets, where relevant.**

We assess cybersecurity risks through our Information Security Division in accordance with the Cybersecurity Management Guidelines of the Ministry of Economy, Trade and Industry (METI) and the NIST Cybersecurity Framework, and formulate organization-wide 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 manage sensitive AI assets as highly confidential information under technical controls such as access controls, encryption, and audit logging. When using sensitive AI assets provided by other companies or incorporating them into products or services, we comply with the management requirements established by the AI vendors providing those assets.

### **12. 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?**

- Internal vulnerability disclosure receipt and tracking processes
- Investigation of reported or identified risks and vulnerabilities
- Development of patches or other mitigations for confirmed risks or vulnerabilities
- Testing of patches or other mitigations
- Scanning for and mitigation of similar risks or vulnerabilities
- Collaborative vulnerability sharing (e.g., with partners, industry groups)
- Vulnerability disclosure program that is aligned with internationally recognized standards, e.g. ISO/IEC 30111 and 29147.
- Reporting to regulators / authorities when relevant
- Recognition of external finders of confirmed risks or vulnerabilities

**13. What post-deployment practices does your organization use to monitor and respond to vulnerabilities, incidents, emerging risks and misuse of advanced AI systems?**

- Conduct post-deployment monitoring
- Perform root-cause analysis and apply lessons learned to improve future development and/or deployment practices

**14. Do you develop or deploy agentic AI systems (systems that can autonomously plan/act or use tools)?<sup>4</sup>**

Yes

**14.A. If yes, please indicate which categories of controls your organization applies. Please select all.**

- Constraints on the agent's action space (e.g., tool-use permissions, environment sandboxing, allowlists/blocklists of resources, scope limits per task)
- Human oversight or interruptibility mechanisms (approval requirements for sensitive actions, escalation paths, interruption interfaces)
- Monitoring (anomaly detection, multi-agent interaction monitoring)
- Logging of agent actions

**Please provide any further comments and implementation documentation.**

*No answer provided*

## Section 3 – Transparency Reporting on Advanced AI Systems

**15. Does your organization publish public reports and/or technical documentation related to the capabilities, limitations, and domains of appropriate and inappropriate use of advanced AI systems? Are these clear and understandable by non-technical audiences?**

Yes

**15.A. If yes, which of the following information is included in your organization's publicly available documentation for any of your AI products?**

- Risk management policies
- Monitoring agent to agent interactions

**Please provide links or references showing where this information is made public.**

- Guiding Principles for the Ethical Use of AI in Social Innovation Business  
[https://www.hitachi.co.jp/products/it/lumada/about/ai/ldsl/document/ai\\_document\\_jp.pdf](https://www.hitachi.co.jp/products/it/lumada/about/ai/ldsl/document/ai_document_jp.pdf)
- Hitachi's Initiatives for the Ethical Use of AI in Social Innovation Business  
[https://www.hitachi.co.jp/products/it/lumada/about/ai/ldsl/document/ai\\_whitepaper01.pdf](https://www.hitachi.co.jp/products/it/lumada/about/ai/ldsl/document/ai_whitepaper01.pdf)

- Practical Initiatives to Implement AI Ethics Principles in Hitachi's Research & Development Group  
[https://rd.hitachi.co.jp/\\_ct/17728022#c17728022\\_h7](https://rd.hitachi.co.jp/_ct/17728022#c17728022_h7)
- To What Extent Can AI Be Trusted? AI Ethics and Social Implementation in Hitachi's Research and Development  
[https://rd.hitachi.co.jp/\\_ct/17830060](https://rd.hitachi.co.jp/_ct/17830060)
- How Have AI Ethics Transformed Hitachi's Approach to Research and Development? Researchers Working to Enhance Reliability Through AI  
[https://rd.hitachi.co.jp/\\_ct/17830075](https://rd.hitachi.co.jp/_ct/17830075)
- Hitachi's Guiding Principles for the Ethical Use of AI and Their Implementation in Social Innovation Business  
<https://www.hitachihyeron.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](https://www.hitachi.co.jp/products/it/appsvdiv/service/justware/ai-apfw/index.html?_fsi=QQkQT4S8)
- A Society with AI Develops Through Dialogue and the Sharing of Values Among Diverse People | Co-Creation Forest Webinar No. 5, "AI Governance," Program 3, "AI Trusted by Society"  
[https://linkingsociety.hitachi.co.jp/\\_ct/17498004](https://linkingsociety.hitachi.co.jp/_ct/17498004)
- Promoting the Use of Generative AI Inside and Outside Hitachi Through the New "Generative AI Center," Accelerating Value Creation in the Lumada Business and Improving 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.hitachihyeron.com/jp/archive/2020s/2024/01/16/index.html#toc-1>
- Hitachi Publishes the Book "Practical Textbook on Generative AI"  
[https://digital-highlights.hitachi.co.jp/\\_ct/17686288](https://digital-highlights.hitachi.co.jp/_ct/17686288)
- Key Global Initiative: Use of Generative AI  
<https://www.hitachi.co.jp/about/it/contents/technology-strategy/contents1.html>

**15.B. If yes, how often are such reports usually updated?**

Other (e.g., triggered by material changes, incidents, or new risks)

**16. 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?**

- Protected information sharing
- Direct engagement with AI safety institutes, civil society or academic partners

**Describe how evaluation results are shared and with which stakeholders (e.g. other organizations, governments, civil society and academia, etc). You may also indicate whether non-technical or plain-language summaries are provided for broader audiences.**

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 the Hitachi AI Advisory Board, which is composed of external experts, as well as other specialists.

Additionally, we participate in ISO/IEC JTC 1/SC 42 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.

**17. Does your organization publish public privacy policies addressing the use of personal data, user prompts, and/or the outputs of advanced AI systems?**

Yes

**If yes, provide links to privacy policies or notices covering personal data, user inputs, and AI-generated outputs.**

We disclose our privacy policy as follows.

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

The Guiding Principles for the Ethical Use of AI also stipulate, as an item to be addressed, the appropriate handling of personal information and the protection of rights, including privacy, with respect to input data used for the training, evaluation, and operation of AI, as well as data output by AI.

**18. How 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?**

- Not applicable

**Please provide links to any publicly available information.**

-

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

Yes

**If yes, describe additional method(s), including where relevant transparency measures related to AI agents**

We promote research and development of explainable AI (XAI) as part of our efforts 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%AAI](https://rd.hitachi.co.jp/_tags/%E8%AA%AC%E6%98%8E%E5%8F%AF%E8%83%BD%E3%81%AAI)

For generative AI agents, we are developing technologies that support human verification of the accuracy of responses by presenting, for each element of a response, the specific passages directly referenced as supporting evidence.

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

**20. Does your organization publish statistical insights derived from anonymised, aggregated usage data to advance understanding of AI adoption and its economic impacts in areas including**

occupations and tasks by region, sector and country? To this end, do you contribute to international cooperation efforts?

No

**Please estimate the total staff time and annual expenses associated with conducting compliance activities and include any further comments and implementation documentation (optional).**

AI-related compliance activities cover a broad range of areas and require diverse expertise not only in AI-related technologies and laws, but also in privacy, copyright, human rights, and other fields. We therefore develop the necessary human resources based on a multidisciplinary approach to AI governance. In other words, personnel in each area of expertise examine matters that should be considered in their respective work from an AI perspective, bring together their findings, and engage in AI-related compliance activities through a multifaceted, cross-functional framework. Because these activities are carried out not only by dedicated organizations but also as part of the regular duties of each organization, it is difficult to accurately calculate the total person-hours and annual costs associated with the compliance activities. However, hundreds of personnel are involved in these activities.

## Section 4 – Organizational Governance, Incident Management, and Transparency

**21. How has AI risk management been embedded in your organization's governance framework?**

- Executive oversight
- Senior management oversight
- Independent advisory body
- Governance council or committee

**Please describe how AI risk management is integrated into governance structures, including: which roles or bodies have advisory versus decision-making authority; which roles or bodies have final accountability for AI risk-related decisions; and how often policies are reviewed or updated (e.g., following major incidents, regulatory changes, or material updates to systems).**

At the company-wide level, the AI Governance Supervisory Committee, established under the CLO, develops and reviews policies, rules, and guidelines for AI governance and oversees AI governance activities across the Hitachi Group.

AI Governance Committees established in each sector, business unit, and Group company conduct risk assessments and provide advice on individual projects, and propose risk mitigation measures as necessary. These AI Governance Committees have advisory authority with respect to AI risks.

By contrast, final decision-making authority regarding whether to proceed with individual AI projects and whether to accept the associated risks rests with the head of the business division or organization responsible for the relevant project. The same person also bears final accountability for AI risk-related decisions concerning the project.

Policies, rules, and guidelines for AI governance are reviewed and updated as necessary in light of advances in AI technology, changes in AI risks, developments in laws and regulations in Japan and

other countries, and lessons learned from incidents.

**22. Are relevant staff trained on your organization's governance policies and risk management practices?**

Yes

**If yes, describe training methods (e.g., onboarding modules, regular sessions, and confidential channels to report concerns).**

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 more practical training through case studies and other methods.

We also operate a community for relevant stakeholders with the aim of accumulating knowledge on AI governance and enhancing practical capabilities. The community shares domestic and international policy and regulatory developments, AI risk cases, technological developments, and practical insights from within and outside the company, thereby promoting continuous learning and mutual knowledge sharing.

**23. Are steps taken to address reported incidents documented and maintained internally?**

Yes

**If yes, describe how incident records are kept and used for follow-up (e.g., incident logs, root-cause analysis reports, internal tracking systems). You may indicate whether and how you define escalation thresholds for when incidents must be surfaced to senior leadership or external parties.**

We document, maintain, and update response rules and related procedures for each type of incident. The company-wide AI governance organization reviews them as necessary in light of changes in AI technologies and risks, as well as developments in domestic and international policies and laws and regulations.

Escalation thresholds have been established based on the severity of the incident, the scope of its impact, legal and contractual requirements, and other relevant factors. Incidents that meet specified thresholds are reported to and shared with the relevant divisions and, where necessary, the appropriate levels of management.

For AI-related incidents determined to be serious, we coordinate with the relevant divisions and, where necessary, share information with Group companies and overseas sites. We also report to and take appropriate action with relevant external stakeholders where necessary, in accordance with applicable laws and regulations, contractual obligations, and other requirements.

Specific escalation thresholds, reporting procedures, and details of internal management processes are not publicly disclosed.

**24. Does your organization share research and best practices on addressing or managing risk?**

Yes

**If yes, describe how research or good practices are shared (e.g., publications, conferences, collaborations with other sectors).**

We share research and best practices on addressing and managing AI risks in the following ways. Internally, we develop guidelines and practical materials on the use of AI that set out key risks associated with AI use, recommended measures, and key considerations. These materials are continuously shared through employee portals and training programs, using specific cases as well. English versions of some of these materials have also been prepared and distributed to Group companies and relevant stakeholders in Japan and other countries.

Externally, we disclose our initiatives relating to the Guiding Principles for the Ethical Use of AI and AI governance on our website as follows.

- 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\\_jp.pdf](https://www.hitachi.co.jp/products/it/lumada/about/ai/ldsl/document/ai_document_jp.pdf)
- Hitachi's Initiatives for the Ethical Use of AI in Social Innovation Business  
[https://www.hitachi.co.jp/products/it/lumada/about/ai/ldsl/document/ai\\_whitepaper01.pdf](https://www.hitachi.co.jp/products/it/lumada/about/ai/ldsl/document/ai_whitepaper01.pdf)
- Practical Initiatives to Implement AI Ethics Principles in Hitachi's Research & Development Group  
[https://rd.hitachi.co.jp/\\_ct/17728022#c17728022\\_h7](https://rd.hitachi.co.jp/_ct/17728022#c17728022_h7)
- To What Extent Can AI Be Trusted? How Hitachi Research and Development Group Is Tackling the Issue of AI Ethics and Social Implementation  
[https://rd.hitachi.co.jp/\\_ct/17830060](https://rd.hitachi.co.jp/_ct/17830060)
- How Have AI Ethics Transformed Hitachi's Approach to Research and Development? Meet the Researchers Working to Enhance Reliability Through AI  
[https://rd.hitachi.co.jp/\\_ct/17830075](https://rd.hitachi.co.jp/_ct/17830075)
- Hitachi's Principles Guiding the Ethical Use of AI and Their Implementation for Social Innovation Business  
<https://www.hitachihyeron.com/jp/archive/2020s/2021/sp/sp01/index.html#toc-12>
- "A Society with AI Develops as Diverse People Engage in Dialogue and Share Values" | Co-Creation Forest Webinar #5, "AI Governance," Program 3, "AI Trusted by Society"  
[https://linkingsociety.hitachi.co.jp/\\_ct/17498004](https://linkingsociety.hitachi.co.jp/_ct/17498004)

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

Yes

**If yes, please indicate which standards or frameworks apply:**

- Other international or regional standards (please specify): AIST's "Machine Learning Quality Management Guidelines" and METI and MIC's "AI Business Guidelines"

**If yes, describe which international technical standards or best practices your organization uses.**

We refer to ISO/IEC 27001, ISO/IEC 42001, the NIST AI Risk Management Framework (AI RMF), and other standards and frameworks. However, we do not rely solely on any particular standard or framework.

Any further comments and implementation documentation.

*No answer provided*

## Section 5 – Content Authentication & Provenance Mechanisms

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

- User disclosure (e.g., text notice, verbal statement, or system label)
- Persistent visual or audio indicators (e.g., watermark, icon, chime, banner) during AI interaction
- Distinct branding or labeling for AI-generated content (e.g., “AI-generated” tag, metadata label)
- Documentation or FAQs explaining when and how AI is used in the system
- Disclosure to end-users when an AI system is acting autonomously or executing actions on their behalf (e.g., making purchases, sending communications, modifying files)

**Provide any additional details or context on how or when users are informed they are interacting with advanced AI systems (e.g. if such mechanisms are planned, in testing, or applied only to certain products).**

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. Furthermore, in response to the implementation of the transparency obligations under Article 50 of the EU AI Act, key considerations regarding the transparency obligations have been included in the Guidelines on Internal Use of Generative AI, and are also scheduled to be included in the company-wide e-learning materials on AI ethics.

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/en](https://www.jstage.jst.go.jp/article/pjsai/JSAI2024/0/JSAI2024_4Xin281/_article/-char/en)

**27. Does your organization use content provenance detection, labeling or watermarking mechanisms that enable users to identify content generated by advanced AI systems?**

Yes

**27.A. If yes, how?**

- Manifest disclosure (e.g., visible label)
- Detection tool(s) for hidden provenance signals (e.g., watermark or metadata checks)

**27.B. If yes, does your organization use international technical standards or best practices when developing or implementing content provenance?**

Yes

**27.C. Does your organization collaborate on or invest in research to advance content authentication and provenance?**

Yes

**Describe how provenance mechanisms are implemented and which standards or tools are used, and any collaborations or investments to improve content authentication (e.g., C2PA credentials, joint R&D projects, open-source initiatives, standard-setting groups). You may also indicate whether these mechanisms are tested for robustness against removal, tampering, or circumvention.**

We are 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/en](https://www.jstage.jst.go.jp/article/pjsai/JSAI2024/0/JSAI2024_4Xin281/_article/-char/en)

**Any further comments and implementation documentation.**

*No answer provided*

## Section 6 – Research & Investment to Advance AI Safety & Mitigate Societal Risks

**28. Does your organization advance research and investment in the following areas?**

- Security
- Safety
- Data quality and representativeness
- Information integrity
- System performance and outcomes (e.g. accuracy, consistency)
- Explainability or interpretability
- Transparency
- Robustness or reliability

**If yes, please describe key research or investment initiatives in the areas selected above (e.g., dedicated safety teams, funding of research programs, contributions to open-source safety tools, publishing research, partnerships and collaborative research initiatives). You may also indicate whether related outputs are publicly shared, shared with partners or within consortia or not publicly disclosed.**

We promote the following research and investment initiatives to enhance AI safety and mitigate societal risks.

[Participation in Standardization and Policymaking]

- Submission of comments on NIST's "U.S. Leadership in AI: A Plan for Federal Engagement in Developing Technical Standards and Related Tools"
- Participation in the standardization of ISO/IEC 42001 (AI management systems)
- Proposal activities for ISO/IEC TS 22440-1 (Functional safety and AI systems)
- Participation as a member of the Ministry of Internal Affairs and Communications' "Conference toward AI Network Society"

- Participation in the Cabinet Secretariat's "Digital Administrative and Fiscal Reform Council" and "Study Group on Data Utilization Frameworks and Systems"

[Research and Development and Knowledge Sharing]

- Ongoing research and development to implement the Guiding Principles for the Ethical Use of AI  
[https://rd.hitachi.co.jp/\\_ct/17728022#c17728022\\_h7](https://rd.hitachi.co.jp/_ct/17728022#c17728022_h7)  
[https://rd.hitachi.co.jp/\\_ct/17830060](https://rd.hitachi.co.jp/_ct/17830060)  
[https://rd.hitachi.co.jp/\\_ct/17830075](https://rd.hitachi.co.jp/_ct/17830075)
- Publication of research findings on AI governance and trust  
<https://www.hitachihyoron.com/jp/archive/2020s/2021/sp/sp02/index.html>
- Promotion of dialogue with society through the Co-Creation Forest Webinar and other initiatives  
[https://linkingsociety.hitachi.co.jp/\\_ct/17498004](https://linkingsociety.hitachi.co.jp/_ct/17498004)

[Specialized Organizations and Frameworks]

- Establishment of a company-wide governance framework by the AI Governance Supervisory Committee
- The Hitachi AI Advisory Board, comprising invited external advisors
- Practical AI ethics initiatives within the Research & Development Group

These outputs are made publicly available in the form of academic papers, industry standards, and policy recommendations, thereby contributing to the enhancement of AI safety across society.

## 29. Does your organization pursue research or investment related to socio-economic or environmental considerations in the development and deployment of AI?

- Socio-economic considerations (e.g., AI adoption and productivity, workforce transitions, skills development, access to AI tools)
- Environmental considerations (e.g., energy use, computational efficiency, resource use)

**If yes, please provide examples of relevant research or investment initiatives for the areas selected above.**

We promote a wide range of investment and research and development initiatives addressing both "socio-economic considerations," such as developing generative AI talent and establishing AI ethics, and "environmental considerations," such as reducing the power consumption of AI processing and improving the energy efficiency of data centers.

### ■ Socio-economic Considerations (AI Adoption, Skills Development, and Governance)

- Large-scale AI talent development and reskilling: With the internal CoE organization, the "Generative AI Center," serving as the hub, we are promoting the development of tens of thousands of AI professionals globally, as well as employees' acquisition and application of AI skills in their work through company-wide workshops (reskilling).  
<https://www.hitachi.com/ja-jp/press/articles/2023/05/0515/>
- Skills transfer and support for workforce transition: We are developing a work support platform that converts the know-how (tacit knowledge) of experienced workers in social infrastructure and manufacturing workplaces into explicit knowledge through domain-specific AI agents and other means, enabling even workers who are still acquiring technical skills to make highly accurate

decisions promptly.

[https://www.hitachi.co.jp/products/it/CloudDX/generative\\_ai/index.html](https://www.hitachi.co.jp/products/it/CloudDX/generative_ai/index.html)

- Hitachi's Guiding Principles for the Ethical Use of AI and governance practices: We have introduced Hitachi's Guiding Principles for the Ethical Use of AI, which stipulate respect for human dignity, the prevention of unfair bias, and the protection of privacy, and have established a governance framework for risk assessments at the research and development stages and for the societal implementation of "responsible AI."

[https://rd.hitachi.co.jp/\\_ct/17728022](https://rd.hitachi.co.jp/_ct/17728022)

- "AI Debriefing Technology" to improve the skills of frontline workers and promote knowledge sharing: This technology enables frontline workers to reflect on the reasons for their decisions, including why they selected a particular response, through dialogue with multiple AI systems, and to organize and articulate the basis for those decisions. While strengthening workers' autonomy and ability to apply their skills (skills development), the technology also transforms the personal experience of skilled workers into knowledge that can be utilized across the organization.

[https://rd.hitachi.co.jp/\\_ct/17840270](https://rd.hitachi.co.jp/_ct/17840270)

#### ■Environmental Considerations (Computational Efficiency, Energy Conservation, and Resource Use)

- Publication of a joint report presenting the opportunities and risks that AI poses to the transition to a sustainable society: In collaboration with Columbia University, we examined how AI may affect the transition to a sustainable society across five areas: the global environment, energy systems, industry and labor, finance, and democracy and societal resilience, and analyzed the associated opportunities and risks. Based on this analysis, the report proposes a governance framework for harnessing the benefits of AI while mitigating its risks, as well as a roadmap for implementing that framework.

[https://rd.hitachi.co.jp/\\_ct/17843022](https://rd.hitachi.co.jp/_ct/17843022)

- Virtual power plant control technology capable of reflecting the diverse values of individual communities in real time: In collaboration with Kyoto University, we developed a new system control technology that enables stable operation of virtual power plants (VPPs) while reflecting, in real time, the diverse values of individual communities, such as cities and regions. In addition to conventional fixed indicators, such as maximizing economic efficiency, this technology can flexibly incorporate context-dependent and changing values, such as CO2 reduction and convenience, into system control.

[https://rd.hitachi.co.jp/\\_ct/17793817](https://rd.hitachi.co.jp/_ct/17793817)

- "Decarbonization Scenario Simulator" to support regional environmental planning and policymaking: We are developing and demonstrating a "Decarbonization Scenario Simulator" that applies an AI simulation technology developed by the Hitachi Kyoto University Laboratory in 2017 to municipal decarbonization. The technology generates possible future scenarios and the "branch points" that emerge as those scenarios are realized. By using AI to simulate key "branch points" when backcasting from a desired future, this AI system visualizes regional challenges and supports the development of decarbonization scenarios while enabling users to envision how the region may look in the near future.

[https://rd.hitachi.co.jp/\\_ct/17685363](https://rd.hitachi.co.jp/_ct/17685363)

**Any further comments and implementation documentation.**

*No answer provided*

## Section 7 – Advancing Human and Global Interests

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

Yes

**If yes, describe any digital literacy or training initiatives your organization supports (e.g., user education, public AI-education programs, school or university partnerships, community workshops). You may also indicate whether these efforts include or prioritise underserved or underrepresented communities.**

We provide learning programs for all employees on AI fundamentals, responsible AI utilization, and AI risk management. These programs offer opportunities to learn about the basic mechanisms and applications of AI, as well as risks associated with AI use, ethical considerations, and appropriate usage methods. These programs are developed by Hitachi or other LMS vendors.

Additionally, for personnel involved in AI governance, we implement more specialized education on AI ethical issues and AI risk assessment and management.

Furthermore, through collaboration with universities, we conduct public symposiums and knowledge-sharing initiatives on the potential and risks of generative AI, human-AI relationships, AI governance, and other topics, working to promote understanding of AI.

We support digital literacy, education, and training through the following initiatives:

- AI ethics training for all employees and internal awareness-raising activities, as well as holding discussion sessions for some employees on global AI ethics trends and research, internal governance, regulations, and other topics

[https://rd.hitachi.co.jp/\\_ct/17728022](https://rd.hitachi.co.jp/_ct/17728022)

[https://rd.hitachi.co.jp/\\_ct/17830060](https://rd.hitachi.co.jp/_ct/17830060)

[https://rd.hitachi.co.jp/\\_ct/17830075](https://rd.hitachi.co.jp/_ct/17830075)

**31. Does your organization pursue AI projects that aim to deliver societal or public benefits, including through partnerships with civil society and community groups?**

Yes

**If yes, please provide examples, including any collaborations with civil society, community groups, or efforts to address broader societal challenges.**

Through our Social Innovation Business, which starts from addressing social challenges, we work to promote benefits for humanity and the planet.

One specific example is the development of “Policy Recommendation AI” in collaboration with Kyoto University. Based on social models, this AI illustrates branching structures and factors among various possible future scenarios, and has been used to analyze Japan’s sustainability in the year 2050. We are also working to enhance sustainability by utilizing AI in the maintenance and operation of social infrastructure. Across business domains such as Energy, Mobility, and Connected Industries, we enable AI-driven prediction, decision-making, optimization, and support for judgment, with the aim of building a trusted technological foundation.

Furthermore, through the development of metaverse, AI, and robotics technologies that support frontline workers, as well as research and development that considers AI transparency and fairness, we contribute to realizing a new society in which AI and humans co-evolve.

As part of our approach that starts from addressing social challenges, these initiatives contribute to resolving challenges faced by civil society and communities.

**Any further comments and implementation documentation.**

*No answer provided*