Source: https://transparency.oecd.ai/reports/634d97e0-6efe-48a1-9183-6313a1b233a2

Report

Organization: **Rakuten Group, Inc.**

https://global.rakuten.com/corp/innovation/

Publication date: Apr 22, 2025, 04:02 PM PDT

Reporting period: 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?

We define and classify risk categories based on Japanese and international technical standards and best practices (e.g., the US National Institute of Standards and Technology (NIST) Risk Management Framework, the OECD AI Principles, Japan's AI Guidelines for Business, and, where applicable, regulations like the EU AI Act), and we evaluate risks based on these categories.

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

The Rakuten Group has established a specialized department in charge of AI safety and security. There are systems in place to proactively develop safe services while eliminating vulnerabilities (information security flaws) and risks by ensuring thorough security and privacy education for developers, implementing security and privacy reviews during the software development process, and conducting inspections for vulnerabilities and risks. Our efforts to prevent AI safety and security incidents also include monitoring illegal access, and surveying and responding to information security and privacy flaws.

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?

The department in charge of AI safety and security conducts general security and safety-related tests to evaluate the suitability of models and systems. In addition, the group's quality control department adds checks to reduce AI risks from the perspective of shipment checks and conducts the necessary verification.

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?

Rakuten Group operates a wide range of businesses both domestically and internationally. As our business expands, we encounter numerous potential risks that could escalate into significant issues. We define risk as "uncertainty that may affect the achievement of management objectives" and we have implemented Enterprise Risk Management (ERM) to enhance the likelihood of achieving these objectives. We have also been operating a Vulnerability Disclosure Program since 2023.

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?

The Rakuten Group has established a Groupwide Computer Security Incident Response Team (CSIRT) to cooperate with external stakeholders such as relevant ministries, organizations specialized in combatting cybercrime, and other security companies, and we are strengthening our cooperation with organizations such as the police and other administrative and investigative agencies, Forum for Incident Response and Security Teams (FIRST), and the Nippon CSIRT Association.

We are committed not only to maintaining our own security but also improving information security for society as a whole. We are also promoting information exchange with the AI Safety Institute (AISI) in Japan and the AI Governance Association in Japan regarding AI-specific risk events.

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?

At Rakuten Group, we are committed to risk management to achieve sustainable development amidst rapid changes in the business and social environment. Our risk management system is built on three key pillars: Enterprise Risk Management (ERM), Incident Management, and Business Continuity Management (BCM).

Under our Group-wide regulations on risk management, we have developed a system that follows a plan–do–check–act (PDCA) cycle for identifying risks, formulating and implementing countermeasures based on their significance, and monitoring the results.

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?

Rakuten Group is a member of the AI Governance Association and actively exchanges opinions with related stakeholders across sectors.

Section 2 - Risk management and information security

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

At Rakuten, when we utilize external generative AI (Gen AI) models, we only adopt external models that satisfy our security requirements. We also conduct security testing before releasing AI systems. Furthermore, we develop and apply guardrail functions to eliminate personal data and harmful information from input into / output from Gen AI models.

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

Based on our test results, we are able to assess the risk by understanding its impact and severity, as well as identifying the necessary countermeasures to mitigate the risk to an acceptable level.

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

All tests are run in a staging environment, not a production environment.

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

At Rakuten, we detect and remove personal data and harmful content, as well as sexual and violent content when we gather training data for Gen AI models. We also conduct bias analyses to maintain equitability and fairness.

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

We require that systems released by the Rakuten Group as a provider be reviewed by the departments in charge of intellectual property before release.

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

We require that systems released by the Rakuten Group as a provider be reviewed by a privacy specialist and the departments in charge of privacy and legal compliance before release. For more details, please refer to the following link. https://global.rakuten.com/corp/privacy/bcr/

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?**

When developing and operating AI models and services, we consider attacks on AI systems, such as data poisoning and prompt injection in Gen AI, and take measures such as risk verification, the introduction of protection functions, and appropriate access control. In addition, the department in charge of AI safety and security conducts general security and safety-related tests to evaluate the suitability of models and systems.

Cybersecurity risks are assessed based on industry standards, guidelines, and frameworks, such as the NIST SP-800 series, Open Worldwide Application Security Project (OWASP) Top 10 for LLM, and MITRE Adversarial Threat Landscape for Artificial-Intelligence Systems (ATLAS). In addition to implementing all general cybersecurity policies, we also establish AI-specific policies to ensure secure development and operations.

In the department that oversees cyber security, we manage vulnerabilities related to the services provided by the Rakuten Group. All vulnerabilities detected during development, testing and operation, as well as vulnerability information received from other stakeholders, are registered in our vulnerability management database and managed accordingly.

The Rakuten Group is registered as a product developer in the Early Warning Partnership operated by the Japan Computer Emergency Response Team / Coordination Center (JPCERT/CC), which is operated by the Information-Technology Promotion Agency, Japan (IPA), and we work closely with them.

At a minimum, reviews are conducted once a year. In addition to the regular review process, we incorporate updates based on lessons learned from our incident response activities and make

adjustments to our security testing methodologies as new techniques are disclosed.

We use a cloud access security broker (CASB) to monitor for threats caused by internal misconduct or human error. In addition, we have set up the Rakuten Hotline, which is an internal and external whistleblowing hotline that all Rakuten Group employees can use to consult or report on any behavior that violates or may violate laws, corporate ethics, or internal regulations, regardless of whether they are full-time, contract, part-time, or other type of employee.

In operating this hotline, and through similar whistleblowing protections in other jurisdictions, we protect the confidentiality of those who consult or report to the hotline, and prohibit and protect against any disadvantageous treatment of those who do so in accordance with applicable laws.

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

We monitor our services, with a focus on the department that handles AI safety and security. We also continuously collect information on incidents and new risks.

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?
- ii. How are new significant releases reflected in such reports?
- iii. Which of the following information is included in your organization's publicly available documentation: details and results of the evaluations conducted for potential safety, security, and societal risks including risks to the enjoyment of human rights; assessments of the model's or system's effects and risks to safety and society (such as those related to harmful bias, discrimination, threats to protection of privacy or personal data, fairness); results of red-teaming or other testing conducted to evaluate the model's/system's fitness for moving beyond the development stage; capacities of a model/system and significant limitations in performance with implications for appropriate use domains; other technical documentation and instructions for use if relevant.

Rakuten plans to release the relevant documents to comply with applicable laws and regulations.

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?

N/A

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

The Rakuten Group discloses its privacy policy for systems it releases as a provider.

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?

Rakuten makes timely disclosures of any data sources that must be disclosed from a legal or regulatory standpoint, except to the extent that an exception applies, including because such information is proprietary information or a trade secret.

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

N/A

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?

At Rakuten, our approach to managing and supervising AI involves several key efforts: AI governance, AI research & development, AI security & safety, public relations & public policy, and data utilization. Within such programs, we actively engage in discussions that cover everything from training our employees to the development and implementation of AI, focusing on identifying challenges and shaping future policies.

The outcomes of these discussions, including proposals and decisions, are reported to the AI & Data Committee. Moreover, key decisions and challenges are regularly reported to the Board of Directors and Corporate Management Meetings, ensuring transparency and accountability in our AI governance. We are also conducting risk assessments in collaboration with experts in each risk category. We have completed risk assessments for 48 use cases from January to March 2025.

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

We provide annual training on the safe use of AI for all our employees. In addition, relevant staff members are also invited to participate in a specialized program dedicated to AI risk management. In 2024, we undertook training for AI auditors provided by ISACA.
https://www.isaca.org/credentialing/ai-audit

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

It is available on our corporate website.

[Risk Management: https://global.rakuten.com/corp/sustainability/risk/]

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

In the event of an incident, we have systems and reporting procedures in place at the Rakuten Group level for implementing measures that minimize the impact on various stakeholders by promptly identifying, assessing, and responding to the incident.

Specifically, the type of incident and the degree of impact - such as financial losses, damage to users, and impact on business continuity and reputation - are evaluated, and responses are defined accordingly. Based on the information collected, we work to prevent the recurrence of incidents by investigating and examining the causes, planning and implementing recurrence prevention measures, and monitoring their effectiveness. Documentation and recordkeeping practices in those regards are equivalent to industry norms and comply with all applicable legal requirements.

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

The Rakuten Group has established a Groupwide Computer Security Incident Response Team (CSIRT) to cooperate with external stakeholders such as relevant ministries, organizations specialized in combatting cybercrime, and other security companies, and we are strengthening our cooperation with organizations such as the police and other administrative and investigative agencies, Forum for Incident Response and Security Teams (FIRST), and the Nippon CSIRT Association.

We are committed not only to maintaining our own security but also improving information security for society as a whole. We are also promoting information exchange with the AI Safety Institute (AISI) in Japan and the AI Governance Association in Japan regarding AI-specific risk events.

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?

We share and disclose appropriate information through a cross-departmental Computer Security Incident Response Team (CSIRT) that is responsible for liaising with external organizations such as related government ministries, specialist cybercrime response organizations, and other companies.

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

We share this information through industry papers and company conferences.

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

Yes. We use international technical standards and best practices (e.g., the NIST Risk Management Framework) for our AI risk management.

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?

N/A

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?

No

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?

N/A

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

N/A

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?

Rakuten is a Corporate Supporter of the OWASP Foundation, and the OWASP deliverables have greatly contributed to making our AI services more robust.

https://owasp.org/supporters/list

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

N/A

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 are committed to leveraging AI to maximize socio-economic and environmental benefits. Through our “AI-nization” initiative, we have integrated AI across our ecosystem to drive greater productivity for both our businesses and clients.

Rakuten AI 7B (Large Language Model): Rakuten received the Open Innovation Field IT Award from Japan Institute of Information Technology for leveraging years of experience and data from e-commerce-specific text processing and a proprietary tokenizer. The lightweight architecture of Rakuten AI 7B was also noted for addressing recent concerns about the significant energy consumption of AI operations.

Rakuten AI 2.0: A large language model and small language model optimized for the Japanese language designed to empower businesses and professionals by enabling cost-effective AI applications. By releasing these models to the open-source community, we are contributing to the broader AI community, accelerating innovation and democratizing access to advanced AI technologies.

We also focus on sustainability by using AI to optimize logistics and reduce carbon emissions in e-commerce and delivery operations. AI-powered route optimization and inventory management systems help minimize waste and energy consumption.

Rakuten Group has published our initiatives regarding AI for 2024 here:

https://rakuten.today/blog/rakuten-ai-in-2024-a-year-of-empowerment.html

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 provide training programs and resources for Rakuten's online marketplace (e.g., Rakuten Ichiba in Japan) merchants through Rakuten AI University to help them understand and effectively utilize AI technologies. We also provide workshops, webinars, and seminars on AI and digital transformation topics.

These events are designed to educate participants on the practical applications of AI, ethical considerations, and the societal impacts of advanced AI systems. By engaging with a broad audience, we aim to empower all sizes of businesses across Japan to maximize the advantages AI offers and promote responsible usage.

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.

Promoting Sustainable Consumption and Production (Sustainable Development Goal (SDG) 12): We use AI to optimize logistics and supply chain operations in the Rakuten Ecosystem. AI-powered route optimization and inventory management systems reduce waste, minimize energy consumption, and lower carbon emissions, contributing to more sustainable consumption and production patterns.

Fostering Quality Education (SDG 4): Rakuten supports AI education and digital literacy through collaborations with universities, open-source AI model releases, and training programs.

Contributing to Industry, Innovation, and Infrastructure (SDG 9): Our release of Rakuten AI for Business and the development of Rakuten AI 2.0 and other advanced AI models support the growth of Japan’s AI industry. By releasing these models to the open source community, we are fostering innovation, collaboration, and the development of AI technologies tailored to Japan's needs.

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.

Open-Source AI Contributions: We have released our advanced AI models to the open-source community. This initiative enables public and private organizations, researchers, and community groups to access cutting-edge AI technologies for free, empowering them to develop solutions tailored to solving challenges we all face.