Infosys Limited: G7 Hiroshima AI Process (HAIP) Transparency Report

https://www.infosys.com

Publication date: Jan 8, 2026, 07:02 AM PST

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?

At Infosys, AI systems can be classified into four risk levels—Prohibited, High, Limited, and Minimal.

Prohibited AI Use Cases: AI systems that pose an unacceptable risk to fundamental rights and safety are prohibited.

High-Risk AI Use Cases: AI systems that significantly impact fundamental rights and safety but are not outright prohibited are considered high-risk. They require stringent obligations before they can be deployed.

Limited Risk AI Use Cases: AI systems in the limited risk category must comply with basic transparency requirements, such as informing users that they are interacting with an AI system.

Minimal Risk AI Use Cases: Minimal risk AI systems pose little to no risk to users' rights or safety and are therefore subject to the least stringent regulatory requirements.

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

Infosys Responsible AI Office continuously does a market scan for AI risks, vulnerabilities, incidents and risks emerging due to misuse of technology. These vulnerabilities are then assessed for exposure in various projects and also identifies fixes to mitigate the vulnerabilities. These are then cascaded to all the project teams that are exposed for such vulnerabilities. Also, all AI Use cases must mandatorily undergo Impact and risk assessment, and the identified risks must be mitigated throughout the AI lifecycle.

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?

Infosys evaluates AI systems through structured adversarial testing, including automated and manual red-teaming, to identify vulnerabilities and ensure ethical, secure, and resilient performance. Findings inform remediation and compliance checks under Responsible AI governance before production deployment, aligning with ISO 42001 and global safety standards.

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?

Yes. Infosys applies both quantitative and qualitative risk metrics during AI evaluations, with caveats wherever applicable. These caveats depend on the type of the project, the applicable jurisdiction and regulations, the AI Use case and the project context. These are documented in governance reviews. Vulnerability and incident reporting channels are accessible to all project stakeholders. Additionally, Infosys encourages responsible disclosure through defined communication channels, ensuring transparency and continuous improvement under its Responsible AI framework subject to client and project confidentiality clauses.

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?

Yes. Infosys engages external independent experts for risk identification, assessment, and evaluation through audits and advisory reviews. The periodic external audits for ISO 42001 and the constant assessment of our best practices with industry leading research analysts help us to identify, evaluate and mitigate the risks. Our strong partner and vendor ecosystem also helps us to proactively identify risks, incidents and vulnerabilities by 3rd parties.

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?

Yes, Infosys both adopts and contributes to international standards for AI risk management.

We are **ISO/IEC 42001 certified**, with processes aligned to global frameworks like the **NIST AI Risk Management Framework**, **EU AI Act**, and other sector-specific guidelines.

We also actively **participate in shaping standards** through consultations and working groups, including **ISO**, **NIST**, **OWASP**, **World Economic Forum (WEF) and Coalition for Content Provenance and Authenticity (C2PA) etc.** ensuring our AI practices reflect global best practices in governance, transparency, and accountability.

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?

Infosys collaborates with clients, industry bodies, and regulatory forums to assess systemic AI risks and implement mitigation measures. Our clients operate across sectors. We participate in global industry bodies like **ISO**, **OWASP** and **Coalition for Content Provenance and Authenticity (C2PA)** to identify the risks. Our regular assessment of AI Use cases through our internal risk management process, helps in identifying the specific systemic risk for a Use case implementation. Our AI risk management process, aligned to global frameworks like NIST AI Risk Management Framework and EU AI Act, helps us to address the identified systemic risks for a Use case implementation.

Section 2 - Risk management and information security

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

Infosys addresses risks across the AI lifecycle through Responsible by Design principles, impact assessments, risk registers, and governance reviews. Measures include bias detection, adversarial testing, privacy safeguards, and continuous monitoring. Findings inform remediation and compliance under ISO 42001 and Responsible AI frameworks, ensuring ethical and secure deployment.

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

Testing measures, including adversarial and stress testing, feed directly into Infosys' AI Life cycle. Identified risks are logged in the risk register, prioritized, and addressed through design adjustments, governance reviews, and compliance checks under Responsible AI process guidelines, ensuring continuous improvement and alignment with ISO 42001 standards

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

Testing occurs in secure, controlled environments during development and pre-deployment phases as per the client and project requirements. The timelines are as per the project schedule and would include isolated infrastructure with strict access controls, encryption, adversarial and stress testing as per the project requirements.

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

Infosys ensures data quality through rigorous validation, provenance checks, and ethical sourcing. Bias mitigation includes diverse dataset curation, fairness audits, and algorithmic reviews during training and collection stages. These measures align with Responsible AI principles and ISO 42001 standards, promoting transparency and reducing harmful bias risks.

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

Infosys safeguards intellectual property through strict compliance with copyright laws, secure data handling, and contractual controls. AI systems are trained on ethically sourced, licensed datasets, and outputs are monitored to prevent infringement. Every AI Use case goes through an internal risk assessment and IP legal check, and copyright violation are integral part of the internal risk assessment process.

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

Our organization protects privacy by restricting the use of personal, sensitive or confidential data and enforcing strict Responsible AI checks with required approvals before any data is used. We document and validate data sources, rights, and security controls to prevent unauthorized disclosure and ensure compliance with privacy requirements throughout the AI lifecycle. Further, every AI Use case goes through an internal risk assessment and Data Privacy check for safeguarding confidential or sensitive data is an integral part of the internal risk assessment process.

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?**

i. Infosys assesses cybersecurity risks through continuous vulnerability scans, penetration testing, and threat modeling for AI systems. Policies enforce encryption, secure coding, and access controls. Governance aligns with ISO 42001 and Responsible AI principles, ensuring resilience against evolving threats and safeguarding advanced AI systems throughout their lifecycle.

ii. Infosys protects critical IP and trade secrets through secure environments with strict access controls, encryption, and role-based permissions. These are managed within client environments for client projects and through licenses for Infosys IP. Proprietary model weights, algorithms, and datasets are stored in isolated infrastructure within client environment with monitoring and compliance checks under Responsible AI governance and ISO 42001 standards, ensuring confidentiality and integrity.

iii. Infosys follows a structured vulnerability management process involving continuous monitoring, risk register updates, and remediation cycles. Identified risks are addressed by first assessing the exposure, finding remediation steps and cascading the impact and remedial measures to the impacted Use case teams.

iv. Infosys reviews security measures regularly through scheduled audits, governance reviews, and compliance checks. Assessments occur at least annually and during major system updates, or on demand depending on the severity of the risk, ensuring alignment with Responsible AI principles, ISO 42001 standards, and evolving cybersecurity best practices for AI systems.

v. Yes. Infosys has an insider threat detection program integrated into its information security framework. It includes continuous monitoring, and strict access controls for sensitive AI assets. These measures align with Responsible AI governance and ISO 42001 standards to prevent unauthorized access and safeguard critical information.

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

To Infosys addresses vulnerabilities, incidents, emerging risks, and misuse across the AI lifecycle through continuous monitoring, risk register updates, and governance reviews. Post-deployment, measures include audits, bias checks, adversarial testing, and tests for identifying model and data drifts.

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.

1. These are updated as per client project requirements for the client Use cases. For internal Use cases, these are updated annually or whenever there is a significant change in the functionality, data or model upgrade.

2. Significant releases are reflected by change in version changes of the technical and / or process documentation and communicated through relevant channels to the appropriate audience.

3. The documentation are available in internal portals and websites and accessible to developers based on their roles and access controls. For client projects, these are subject to client project requirements.

---

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?

The outcome of evaluations of risks and impacts to the client Use case are shared with the relevant stakeholders as per the client contractual agreements.

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

These are subject to client contractual agreements and shared with the relevant stakeholders as per the project requirements.

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?

Yes, subject to client contractual agreements and project requirements.

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

Yes, subject to client contractual agreements and project requirements.

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?

Our organization has a comprehensive ISO 42001 certified AI Governance Program. The program provides end-to-end oversight across the AI lifecycle, covering risk management, accountability, transparency, human oversight, data governance, and regulatory compliance. These are updated annually or when triggered by regulatory changes, emerging risks, audits, or technological advancements, ensuring alignment with ISO 42001 standards and global best practices.

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

Yes. Infosys trains relevant staff through structured programs, including Responsible AI literacy workshops, e-learning modules, and governance briefings. Training covers risk management practices, ethical AI principles, and compliance requirements, ensuring employees understand and apply policies effectively across the AI lifecycle

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

Yes. Infosys communicates its risk management policies to users through Responsible AI documentation, and to public through various public events. These includes publishing ethical guidelines, participating in industry forums, events, awards and sharing governance practices via consortiums and regulatory platforms, ensuring transparency and alignment with ISO 42001 standards

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

Yes, all reported incidents and their corresponding mitigation steps are documented and maintained centrally under the **Responsible AI Office**. Each incident is recorded in a structured manner, followed by root cause analysis and closure tracking to support audit and compliance requirements.

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

Infosys shares information on vulnerabilities to all the relevant teams over email and incidents, AI risks through internal governance forum, the **AI War Room**, OFFERING reports and cross-functional reviews involving Legal, Risk, and Security teams to ensure timely awareness and coordinated action.

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?

Yes, Infosys shares relevant AI incident information with stakeholders such as internal governance bodies, clients, and regulatory partners as appropriate. Public reporting is done only when required under compliance or contractual obligations, ensuring confidentiality and responsible disclosure.

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

We generate market scan reports every month which gives a snapshot of the AI market (especially enterprise AI), including trends, regulatory shifts, risks, opportunities, and how organizations can adopt AI responsibly.

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

Yes. Infosys aligns its Responsible AI framework with global standards, including ISO/IEC 42001, NIST AI Risk Management Framework, EU AI Act, and UNESCO ethical principles, ensuring robust governance, risk mitigation, and compliance across AI lifecycle

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?

Infosys advocates transparency by implementing clear **AI interaction disclosures**. Users are informed when engaging with AI systems through explicit notifications and labeling mechanisms, aligned with Responsible by Design principles for ethical and accountable AI deployment. These include visible disclosures, onboarding messages, and accessible documentation. For sensitive use cases, consent options and ability to engage a human are provided. The implementation of these are per the client and project requirements.

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?

Yes. Infosys advocates content provenance measures like labeling and watermarking for AI-generated outputs, ensuring transparency. These mechanisms follow global best practices and align with standards such as ISO/IEC 42001 and NIST AI RMF for ethical and accountable AI governance. The implementation of these are per the client and project requirements.

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?

Infosys advances research and investment through its Responsible AI Office, focusing on security, fairness, transparency, explainability, and robustness. Initiatives include bias detection, adversarial resilience, and open-source Responsible AI Toolkit, aligned with global standards like ISO 42001 and NIST RMF for trustworthy AI systems. This is accelerated by regular market scan which feeds to the research team.

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

Infosys collaborates with global forums and invests in research to strengthen content authentication and provenance. Initiatives include partnerships, open-source tools, and adherence to guidance provided by Coalition for Content Provenance and Authenticity (C2PA) and ISO/IEC.

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?

Yes. Infosys actively participates in global collaborations and research projects to advance AI safety, security, and trustworthiness. Investments include participation in and contributions to working groups of EU AI Office, NIST, WEF, ISO and UNESCO.

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

Infosys invests in Responsible AI research to reduce socio-economic and environmental risks by promoting fairness, inclusivity, and sustainability. Initiatives include bias mitigation, energy-efficient AI practices, and alignment with global standards like UNESCO recommendations for Ethical AI and sustainable AI development. Infosys invests in sustainable AI research by promoting ethical, efficient, and responsible AI development, including autonomous systems, that are designed to optimize resources and reduce the overall carbon footprint.

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.

Infosys invests in AI research to maximize socio-economic and environmental benefits through initiatives like energy-efficient AI, inclusive solutions, and sustainability-focused projects. Examples include AI for renewable energy optimization, open sourcing the Infosys Responsible AI Toolkit and social impact programs through collaborations with industry forums, academia and startups.

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.

Yes. Infosys supports digital literacy and AI education through initiatives like **Infosys Springboard** and **AI training programs**, helping users understand AI's capabilities, limitations, and impacts. These programs promote responsible AI awareness and ethical use globally.

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.

Yes. Infosys prioritizes AI projects aligned with UN Sustainable Development Goals through its Responsible AI framework, focusing on human-centric solutions. Examples include AI for renewable energy optimization and social impact initiatives promoting inclusivity and sustainability. We embed responsible AI principles such as fairness, transparency, privacy, and human oversight across the full AI lifecycle through our governance framework and Responsible AI Toolkit.

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.

Yes — Infosys collaborates with civil society, academia, and community organizations to develop AI solutions that support the UN SDGs. For example, we co-developed the AI-powered Sustainability Atlas with Economist Impact (Economist Impact and Infosys Launch The Sustainability Atlas to Help Businesses Navigate a Sustainable Future) to advance climate and social decision-making, and through the Infosys Foundation's Aarohan Social Innovation Awards, we support NGOs and social innovators building technology for education, healthcare, and environmental sustainability. These partnerships ensure our AI initiatives address real community needs and contribute to inclusive and sustainable development.