

OECD AI Transparency Report

Organization: EXCELYA (FRA) 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?¹

- Align with specific external frameworks, standards, reports, and/or guidance documents such as taxonomies, classification systems, and/or sets of definitions for risk areas (e.g. the NIST AI Risk Management Framework, the ISO/IEC 42001, the OECD AI Principles, the EU General-Purpose AI Code of Practice, the International AI Safety Report). (please specify): NIST AI Risk Management Framework, OECD AI Principles, the EU General-Purpose AI Code of Practice

AI Governance Evaluator

Response Relevancy

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.

Excelya does not use one fixed numerical threshold to determine when a severe AI risk becomes unreasonable.

We use a risk-based governance framework. This framework includes AI governance, privacy, quality, security, and human oversight requirements.

We assess each AI system based on:

- Its intended use.
- The type of data that it processes.
- Its potential effect on people and the organization.
- Applicable legal and regulatory requirements.
- The effectiveness of available risk controls and mitigations.

The AI Lab performs technical, privacy, quality, and security reviews.

In addition, all significant AI use cases receive an independent review from the Digital Committee before deployment.

The Digital Committee operates separately from the AI Lab.

The Committee provides an additional level of governance and oversight.

During the review, the Committee evaluates:

- The business need for the AI use case.
- The risk profile of the AI system.
- Regulatory and ethical considerations.
- Data governance requirements.
- Potential effects on users.
- The adequacy of proposed controls.

The Committee must complete its review before approval for production use.

If the Committee identifies concerns, it can require:

- Additional risk mitigations.
- Additional testing.
- Additional documentation.
- Additional monitoring measures.
- Restrictions on the deployment scope.

The organization must implement these actions before authorization is granted.

Governance responsibilities are divided as follows:

1. The AI Lab designs and validates the solution.
2. Business, Data Protection, and Quality Assurance stakeholders review risks in their areas of responsibility.
3. The Digital Committee independently approves, rejects, or challenges deployment decisions.

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?

- Structured techniques to explore the behavior or performance of a model and/or application, such as model elicitation or bottleneck assessments
- Impact assessments such as personal data protection assessment and fundamental rights impact assessments
- Incident reports, including reports shared by other organizations

- Automated monitoring / remote status reporting (e.g., telemetry pipelines, automated alerts when risk thresholds are breached)
- Specific risk identification practices for agentic AI capabilities, e.g., autonomous action-taking, tool use, multi-step planning (please specify): We develop skills and agentic handover to kill actions that drop under a pre-set score
- Assessment of risks arising from interactions among multiple AI agents
- Multilingual and cross-cultural evaluation, including collaboration with local ecosystems (e.g., national AI institutes, language-specific evaluation consortia) for under-represented languages and cultural contexts (please specify): We operate in Greek, French, and English in medical, legal, and HR contexts. The AI-Lab includes test dataset across the languages to ensure responses are consistent, especially where incorrect grammar can slip in, and review accuracy with native speakers.
- Qualitative assessments (e.g., expert judgment, narrative scoring)
- Internal red teaming³
- External model, system, or application-independent research initiatives

Accuracy

How does your organization monitor the interactions of AI agents under your control with other agents / assess risk from multiple agents interacting?

Our organization deploys AI agents exclusively in the Microsoft ecosystem. This ecosystem includes Microsoft 365 Copilot, Copilot Studio, Azure AI services, and related orchestration services.

We assess the risks of interactions between AI agents. This assessment is part of our AI governance process, architecture review, security review, and production readiness review.

We do not permit unrestricted autonomous interactions between AI agents.

AI agents must operate:

- Within approved workflows.
- Within defined authorization limits.
- Within documented business processes.

We define monitoring requirements during solution design. These requirements remain active during the complete operational life of the solution.

Monitoring activities can include:

- Review of execution logs.
- Review of workflow history records.
- Review of error records.
- Review of conversation logs, when available and enabled.
- Review of test evidence.
- Monitoring of system performance.
- Review of incident management records.

We use this monitoring information to detect issues, verify compliance with requirements, and support continuous improvement of AI solutions.

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.

Our organization uses a defined governance process before deployment of an AI model or system.

The AI Lab designs, develops, tests, and validates the solution.

Business stakeholders, Data Protection, Quality Assurance, and Security representatives review risks in their areas of responsibility.

The Digital Committee performs an independent review before deployment. The Digital Committee has authority to approve, reject, or request additional controls for significant AI use cases.

Before deployment, the project team must provide evidence for:

- Functional performance.
- Security controls.
- Privacy controls.
- System robustness and reliability.
- Compliance with applicable requirements.
- Risk assessment results.
- Test, evaluation, validation, and verification activities.
- Known limitations and residual risks.

Risk evaluation results must be interpreted in the context of the intended use, operating environment, data quality, available controls, and identified limitations. Risk metrics alone do not determine deployment approval.

The following approvals can be required before deployment:

- Business Owner approval.
- Data Protection approval, when applicable.
- Quality Assurance approval, when applicable.
- Digital Committee approval for significant AI use cases.

Before deployment, the organization must implement appropriate safeguards. These safeguards can include:

- Technical and procedural risk mitigations.
- Access controls and authorization limits.
- Human oversight measures.
- Monitoring and logging controls.
- Incident management procedures.
- Usage restrictions and operational guardrails.
- Additional testing and validation activities.

The organization can restrict deployment scope or delay deployment if the identified controls are not sufficient to reduce risk to an acceptable level.

After deployment, the organization continues monitoring activities and reviews the solution regularly to verify compliance, performance, and effectiveness of controls.

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.

a { text-decoration: none; color: #464feb; } tr th, tr td { border: 1px solid #e6e6e6; } tr th { background-color: #f5f5f5; }

After deployment, our organization reviews incident reports from internal and external sources.

We use these reports to identify new risks, recurring issues, and changes in the risk profile of an AI system.

The review team assesses the relevance of each incident to the deployed system and its intended use. Each identified issue receives a risk score. The risk score determines the required response time and corrective actions.

The organization uses defined remediation timelines that range from 2 days to 30 days, based on the assigned risk score.

If the review identifies a risk, the organization updates the risk assessment and determines the required actions.

Corrective actions can include:

- Additional monitoring.
- Additional testing and validation.
- Updates to documentation.
- Changes to controls or mitigations.
- Restrictions on system use.
- Model replacement or retirement when required.

For example, if a model reaches its expiry date or no longer meets governance requirements, the organization reviews the impact and takes the appropriate action before continued use.

The organization records identified issues and tracks corrective actions through its monitoring and incident management processes.

The organization verifies that corrective actions are completed within the required timeline for the assigned risk score.

Monitoring and incident management activities remain active throughout the operational life of the AI system.

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 / as relevant	Not currently accessible
Relevant employees (security, legal/compliance, etc.)	☐		

	Always	Ad hoc / as relevant	Not currently accessible
Public authorities (e.g., public reporting offices for security incidents or data breaches)		☐	
Supervisory authorities (e.g., data protection authorities)		☐	
Industry partners or industry fora			
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			
Independent auditors			
Standards bodies / multi-stakeholder initiatives (OECD.AI, ISO/IEC, IEEE, etc.)			
Specific civil society organizations			
Media / general public (people or communities indirectly impacted)			
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”

Our organization makes reporting mechanisms available according to applicable legal requirements, internal policies, and the risk level of the reported issue.

Employees are trained to report AI-related incidents via email to an internal AI inbox. Suspected data breaches are reported to IT.

How does your organization share relevant information as appropriate?

- Reporting to regulatory authorities as relevant

Please describe what type of information is shared.

Incident report for serious incidents under the AI Act (High-risk AI systems)

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

- Not applicable

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 and uses

If contributes or uses, please select all that apply:

- NIST AI Risk Management Framework (AI RMF)
- Industry-led best practices or codes of conduct

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.)		☐	
Public authorities / regulators / AI safety institutes / other public sector bodies and inter-governmental coordination structures			
Industry partners or industry fora		☐	
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		☐	
Independent auditors			
Standards bodies / multi-stakeholder initiatives (OECD.AI; ISO/IEC; IEEE; AIID etc.)			
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.

We partner with industry groups such as Drug Information Association (DIA) and MAPS (Medical Affairs Professional Society) to share benchmarks, experience and best practices. We are a member of IASEAI, and collaborate with academic centers (The Open University, KCL...) on specific topics (specifically automation bias and HiTL in medical affair processes).

Any further comments and implementation documentation.

NA

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).

Our organization applies risk management during the complete AI system lifecycle.

The AI Lab designs, tests, validates, documents, and monitors AI systems. Business, Quality, Security, and Data Protection stakeholders review risks in their areas of responsibility. The Digital Committee performs an independent review of significant AI use cases before deployment. Significant risks can be escalated to the Digital Committee for additional review and decision.

Our AI systems use Microsoft technologies, including Microsoft 365 Copilot, Copilot Studio, Azure AI services, Power Automate, and Microsoft Graph connectors. We use Microsoft evaluation and observability capabilities to assess quality, grounding, hallucination rates, prompt injection resistance, indirect prompt injection resistance, output safety, content filtering effectiveness, agent tool usage, performance, and reliability. We also use workflow run history, error logs, conversation logs (when enabled), audit records, and operational monitoring data to support governance, traceability, and operational oversight.

Before deployment, each AI system undergoes risk assessment, testing, evaluation, validation, security review, privacy review, and documentation review. Higher-risk systems receive additional controls and oversight. Agentic AI systems receive additional review of autonomy, orchestration logic, authorization limits, tool access, multi-agent interactions, and human oversight mechanisms.

After deployment, the organization monitors evaluation results, system performance, execution logs, workflow histories, error records, conversation logs, and incident reports. Each issue receives a risk score. The risk score determines the escalation path and remediation timeline. Corrective actions must be completed within defined timelines ranging from 2 to 30 days.

The organization reviews its AI governance practices regularly and whenever significant events occur, which in practice has meant every quarter! Risk assessments, evaluation results, and controls are updated as required

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).

Our organization does not develop or train foundation AI models.

We use third-party AI models and apply governance controls to the data, knowledge sources, testing datasets, and evaluation processes used in our AI solutions.

Before use, datasets and knowledge sources undergo review for quality, relevance, accuracy, completeness, and compliance with applicable privacy requirements.

Personal information is removed or anonymized where required. Data used for testing and evaluation is curated and approved before use.

The organization uses a combination of internal datasets and publicly available reference datasets where appropriate. In the clinical research and clinical trial domain, publicly available scientific, regulatory, and industry-standard data sources can be used to support data quality assessment, benchmark development, and evaluation activities.

Domain experts review datasets and evaluation results to verify accuracy, consistency, and suitability for the intended use case.

Data can be labelled manually by subject matter experts or through controlled data-processing pipelines. The organization performs quality checks to verify the correctness of labels and metadata. The organization uses Microsoft evaluation capabilities in Azure AI services and Copilot Studio to assess output quality, grounding, hallucination rates, safety, fairness, reliability, and performance. Test datasets are used to evaluate system behavior under representative use conditions. The organization reviews known limitations of datasets, evaluation datasets, and knowledge sources. Risk assessments consider data coverage, potential bias, domain limitations, and the ability of evaluation results to generalize to real-world use cases. If testing identifies quality, fairness, safety, privacy, or performance concerns, the organization implements corrective actions before deployment. These actions can include dataset updates, additional review by subject matter experts, changes to prompts or guardrails, additional testing, or restrictions on system use. The organization maintains monitoring and periodic review activities to verify that data quality, system performance, and risk controls remain effective throughout the operational life of the AI system.

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

- Restrictions on outbound Internet or API access, including a list of permitted domains or APIs
- Isolation of staging environments (from production / operational environments)
- Rate limits and resource caps
- 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.

This is shared across IT, the AI_Lab, and selected internal Champions.

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:

- Not applicable

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

- Not applicable

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

- Use security controls, including encryption at rest and in transit

- 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.

Our organization applies privacy and data-protection controls throughout the AI system lifecycle.

We use a combination of technical, organizational, and procedural controls to protect data and reduce privacy risks.

Personal information is removed, anonymized, or minimized when required by the intended use case and applicable regulations.

We use role-based access controls to limit access to data, systems, tools, and AI capabilities. Users receive only the access required for their job responsibilities.

We use encryption and secure data-handling practices to protect data during storage and transmission.

We implement guardrails, access restrictions, content controls, and approved skills or tools to prevent unauthorized actions, unauthorized data access, and unintended system behavior.

Before deployment, AI systems undergo privacy, security, and compliance reviews. These reviews verify that data sources, processing activities, access permissions, and controls are appropriate for the intended use case.

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
- 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
- Implement incident response, disaster recovery, or business continuity plans, including escalation procedures for high-severity incidents
- Maintain vulnerability management processes (scanning, channels for external reporting, patching, remediation)
- Integrate cybersecurity AI risks into broader cybersecurity risk management activities
- Regularly review security measures and update them in response to emerging risks, incidents, or major releases
- These responsibilities are delegated contractually to model or infrastructure providers

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
- Conduct secure AI development and/or deployment practices
- Conduct security testing focused on AI-specific risks (e.g., red-teaming, adversarial testing, stress tests)

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

- These responsibilities are delegated contractually to model or infrastructure providers

11.D. Insider risk protection detections

- Conduct pre-employment screening or role-sensitive vetting
- Monitor for unauthorized internal access or suspicious activity
- 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.

Cannot provide response for security reasons!

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?

- Other(s) (please specify): Unable to share

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

- Monitoring and emergency controls specific to agentic AI systems (e.g., ability to interrupt autonomous action sequences, override tool-use permissions, or revoke delegated authority)
- Perform root-cause analysis and apply lessons learned to improve future development and/or deployment practices
- Conduct post-deployment monitoring

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

Yes

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

- 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
- Constraints on the agent's action space (e.g., tool-use permissions, environment sandboxing, allowlists/blocklists of resources, scope limits per task)

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?

No

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?

- Other (specify): Comité Social d'Entreprise

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.

Evaluation results are shared with stakeholders according to the purpose, risk level, and applicable legal or regulatory requirements.

Internally, evaluation results are reviewed by the AI Lab, business owners, Quality Assurance, Information Security, Data Protection, and the Digital Committee when required.

In France, information related to AI deployments is also shared with employee representative bodies when required by applicable legal obligations, including consultation and information processes involving the **Comité Social et Économique (CSE)**.

Evaluation results support deployment decisions, risk assessments, corrective actions, monitoring activities, and governance reviews.

The organization provides both technical and non-technical summaries. Technical reports are used for validation, governance, and audit purposes. Plain-language summaries are used when communicating with business stakeholders, employees, management, and other non-technical audiences.

Selected results, lessons learned, and governance practices can also be shared with customers, partners, professional networks, industry forums, academic collaborators, conference audiences, or regulatory stakeholders when appropriate.

The organization shares information in a manner that protects confidential information, personal data, intellectual property, and applicable regulatory obligations.

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

No

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.

Not applicable

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

No

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).

The organization estimates that AI compliance and governance activities require approximately **0.5 to 1.0 full-time equivalent (FTE)** per year across the AI Lab, Quality, Data Protection, Security, and governance functions.

Activities include:

- Risk assessments.
- AI governance reviews.
- Documentation and record maintenance.
- Testing, evaluation, validation, and verification activities.
- Privacy and security reviews.
- Training and AI literacy activities.
- Monitoring and incident management.
- Regulatory and governance reporting.
- Employee information and consultation activities where required.

Direct annual expenses are primarily associated with governance activities, employee training, testing and evaluation activities, Microsoft AI governance and monitoring capabilities, and participation in professional and

regulatory initiatives.

The organization uses existing Quality Management System (QMS), information security, privacy, and governance processes whenever possible. This approach reduces the need for dedicated compliance infrastructure and supports integration of AI governance into existing operational processes.

Compliance activities are reviewed and updated regularly to address regulatory developments, new AI capabilities, identified risks, incident findings, and evolving business requirements.

Section 4 – Organizational Governance, Incident Management, and Transparency

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

- 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).

Our organization integrates AI risk management into its overall governance framework.

The AI Lab provides technical, operational, and governance expertise. The AI Lab performs risk assessments, testing, validation, documentation, monitoring, and recommendations for AI systems. Business stakeholders, Data Protection representatives, Quality Assurance personnel, Information Security personnel, and subject matter experts provide advisory review within their areas of responsibility.

The Digital Committee performs an independent governance review of significant AI use cases. The Digital Committee has decision-making authority for approval, rejection, escalation, or the application of additional controls before deployment.

Final accountability for AI risk-related decisions remains with the designated business owner and the organization's governance structure. Deployment cannot proceed without the required reviews and approvals.

AI governance policies, risk assessments, controls, and operating procedures are reviewed on a regular basis and whenever significant events occur, in practice this has been every quarter.

Review and update triggers can include:

- Security incidents.
- Privacy incidents.
- Compliance findings.
- Regulatory or legal changes.
- Significant model or system performance issues.
- Material changes to AI system functionality.
- New AI capabilities or agentic features.
- Newly identified risks.
- Model expiry or retirement events.

When a review identifies new or increased risk, the organization updates the risk assessment and determines the required corrective actions. Corrective actions can include additional testing, additional monitoring, updated controls, deployment restrictions, model replacement, or model retirement. For higher-risk AI systems, including agentic AI systems, the organization applies additional governance, review, monitoring, and human oversight requirements before and after deployment.

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).

Our organization provides AI literacy and AI governance training to employees and relevant stakeholders. The objective is to ensure safe, responsible, compliant, and effective use of AI systems. Training programs are aligned with AI governance requirements and AI literacy obligations.

Training methods include:

- Mandatory onboarding and foundational AI training.
- Role-based training modules tailored to business responsibilities.
- Regular awareness and upskilling sessions.
- AI workshops, webinars, and practical learning activities.
- Guidance documents, prompt libraries, and approved-use resources.
- Assessments and certification activities to verify understanding of AI risks and responsibilities.

Training topics include:

- AI governance and legal responsibilities.
- AI risk awareness and escalation procedures.
- Safe prompting practices.
- AI hallucinations and output verification.
- Data sensitivity, confidentiality, and privacy protection.
- Human oversight and accountability.
- Tool scope, limitations, and approved use.

Employees can raise questions, request guidance, or report concerns through established governance, quality, security, data-protection, and incident-management channels. Issues can be escalated to the AI Lab, relevant governance stakeholders, or the Digital Committee, depending on the nature and severity of the concern.

Training content and governance guidance are reviewed and updated when significant regulatory changes, new AI capabilities, policy updates, or new risk scenarios are identified.

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.

AI-related incidents are managed through the organization's Quality Management System (QMS) and follow the same incident-management process used for other quality, compliance, security, and operational incidents.

Incident records are maintained in the QMS and include the incident description, impact assessment, risk score, investigation records, corrective actions, preventive actions, approvals, and closure status. The organization performs investigations and root-cause analysis when required. The organization uses incident records to identify trends, evaluate recurring issues, verify the effectiveness of corrective actions, and improve risk controls.

Each incident receives a risk score. The risk score determines the response priority, escalation requirements, and remediation timeline. Defined remediation timelines range from 2 days to 30 days, depending on the severity of the issue.

Significant incidents are escalated through the organization's governance structure. Escalation can involve business owners, Quality Assurance, Information Security, Data Protection representatives, the AI Lab, the Digital Committee, or senior leadership, depending on the nature and severity of the issue. Where required by applicable legal, regulatory, contractual, or organizational obligations, incidents can also be reported to external authorities or other relevant parties.

The organization tracks all corrective and preventive actions until completion and verifies their effectiveness before incident closure.

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

No

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:

- ISO/IEC standards (e.g., ISO/IEC 27001, ISO/IEC 42001)
- NIST AI Risk Management Framework (AI RMF)
- Industry-led best practices or codes of conduct
- OECD AI Principles
- Other international or regional standards (please specify): EU AI ACT

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

We mostly follow NIST, and ISO, within overarching EU AI Act considerations.

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)
- Other (specify): Bear avatar

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).

Excelya informs users when they interact with AI systems.

All chatbot solutions include a clear disclaimer that identifies the system as an AI-based system and explains that outputs require appropriate human review.

The organization applies transparency measures during solution design and deployment. Users receive information about the purpose, scope, capabilities, and limitations of the AI system when appropriate.

The organization made an early design decision not to use human avatars or AI-generated human images in situations where users could reasonably mistake an AI system for a human.

Instead, the organization uses a clearly identifiable non-human character, the robot teddy bear **BEX**, in interfaces, communications, and training materials where there is a risk of confusion between human and AI-generated content.

This approach helps users distinguish AI systems from human users and supports the organization's transparency, trust, and responsible AI objectives.

The organization reviews transparency measures as part of its AI governance process and updates them when new capabilities, new use cases, or new regulatory requirements are introduced.

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

No

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?

- Other (please specify): Professional judgement, responsibility, and human in the loop

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.

Excelya invests in research and practical initiatives that support responsible use of AI in regulated environments.

A major area of focus is Human-in-the-Loop (HITL), professional judgement, and human oversight in Medical Affairs, Medical Information, pharmacovigilance, regulatory, and compliance workflows. Our work examines what professionals need in order to exercise meaningful judgement when reviewing AI-assisted outputs, including access to evidence, decision authority, escalation pathways, and the ability to challenge or reject AI-generated content.

We are particularly interested in situations where individuals remain accountable for decisions supported by AI. Current work explores how organizations can prevent oversight from becoming a simple approval step and instead support genuine review, critical thinking, and professional responsibility.

Research topics include professional judgement, cognitive workload, automation bias, trust in AI systems, human-AI collaboration, auditability, and governance. The objective is to develop practical approaches that organizations can use to evaluate whether Human-in-the-Loop controls are effective in real-world regulated workflows.

We intend to share selected findings through scientific and professional channels. Planned outputs include a potential manuscript for Therapeutic Innovation & Regulatory Science (TIRS) and conference presentations or posters at events such as AISEAI. The goal is to contribute practical evidence and governance guidance for organizations that deploy AI in regulated environments.

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

- Not applicable

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.

Excelya supports AI literacy, education, and awareness initiatives for employees and external audiences.

Internally, the organization provides AI training programs, AI literacy modules, workshops, awareness sessions, governance guidance, and practical learning resources. These activities help employees understand the capabilities, limitations, risks, and appropriate use of AI systems. Training topics include AI governance, AI risk awareness, data protection, confidentiality, human oversight, output verification, and responsible use of AI.

The organization also supports AI education through internal communities and knowledge-sharing initiatives, including AI Champions and incentivized employee engagement programs. These activities encourage responsible AI adoption, practical learning, and discussion of emerging AI risks and opportunities.

In addition, Exelyates participate in community education activities outside the workplace. These activities include discussions and awareness sessions with parents, teachers, and students in preschool and primary school settings. The objective is to help children and caregivers understand the nature, capabilities, limitations, and responsible use of AI technologies.

These initiatives place particular emphasis on AI literacy, critical thinking, human oversight, privacy awareness, and (in)appropriate reliance on AI-generated content. The objective is to improve understanding of AI among both professional users and the wider community.

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

No

Any further comments and implementation documentation.

No answer provided