

OECD AI Transparency Report

Organization: Criteo (LUX)

Reporting Period: Q3 2026

Published: October 2, 2026

Section 0 – Your role in the AI system lifecycle

Responding as (please select all that apply):

- Model developer (/provider)
- Application developer (/provider)
- Deployer

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

- Maintain an internal taxonomy, classification system, and/or set of risk area definitions
- 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): Criteo does not develop advanced AI systems, meaning systems with general and sophisticated capabilities that may create high risks to safety, security or society. Criteo aligns with the AI Act and its guidelines and the AI PACT.

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.

Criteo does not develop advanced AI systems, so capability-based thresholds are not the relevant framework.

Given the nature of Criteo's services — online advertising based on pseudonymized data — severe-risk scenarios contemplated by capability-threshold frameworks do not generally arise.

Criteo AI products are classified against the risk taxonomy of the EU AI Act (prohibited, high-risk, limited-risk, minimal-risk) and are subject to the relevant graduated obligations tied to the risk classification.

Beyond legal requirements, Criteo may also restrict the use of AI models and systems in specific contexts as per its advertising guidelines: <https://www.criteo.com/advertising-guidelines/>

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

- Adversarial testing to detect technical vulnerabilities
- Publicly available benchmarks and measurement tools
- Internal benchmarks, standard metrics (e.g., frequency, severity scores)
- Specific risk identification practices for vulnerable groups
- Impact assessments such as personal data protection assessment and fundamental rights impact assessments
- Internal model, system, or application-independent research initiatives
- External model, system, or application-independent research initiatives
- Vulnerability reports
- 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): Agentic AI security

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.

The Product and R&D organizations are accountable for the AI products they develop or deploy. AI initiatives are subject to a risk analysis based on the applicable EU AI Act risk category in close coordination with the Privacy and Legal departments. AI products, including third-party AI, must comply with applicable privacy, security, intellectual-property, confidentiality, fairness, safety and transparency requirements

AI models and systems are validated using rigorous methodologies, relevant statistical techniques and appropriate datasets. They are tested, validated, monitored and documented in a controlled environment under varied conditions.

Novel AI tools require prior authorization from the Product, Privacy, Legal and IT departments before live testing with Criteo data can begin.

Required safeguards include appropriate user disclosures, human oversight, filtering or deletion procedures for relevant data, operational monitoring, and the ability to override, modify, isolate or turn

off the AI system when needed. Disclaimers must be validated by the legal department where applicable.

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

We continuously reassess risks by monitoring the scientific literature and working with the relevant regulators.

For instance, we participated to a testing pilot with the French data protection authority to assess privacy risks in AI models (PANAME project). <https://www.cnil.fr/fr/projet-paname-participez-aux-tests-dun-outil-daudit-rgpd-des-modeles-dia>

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.)	☐		
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			

	Always	Ad hoc / as relevant	Not currently accessible
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”

Availability and level of detail depend on the nature and severity of the report, the data involved, the legal obligations, and the need to protect ongoing investigations.

- [Criteo Vulnerability Disclosure Policy](#)
- [Criteo Security Trust Center](#)

How does your organization share relevant information as appropriate?

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

Please describe what type of information is shared.

Shared information depends on the risks, incidents or vulnerabilities. It generally relates to affected products or components, severity and impact, indicators, subject to relevant legal, regulatory and contractual requirements.

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

- Paid rewards (“bug bounty”)
- Vulnerability Disclosure Programs (VDPs) with safe harbor / liability protection
- Research grants, contracts, and hiring red teamers
- Internal incentives and culture

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

<https://www.criteo.com/trust-center/security/vulnerability-disclosure-policy/>
<https://team.inria.fr/fairplay/>

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?

Uses

If contributes or uses, please select all that apply:

- ISO/IEC standards (e.g., ISO/IEC 27001, ISO/IEC 42001, ISO/IEC 23894, ISO/IEC 42005)
- Industry-led best practices or codes of conduct
- Other international or regional standards (please specify): We participated to a testing pilot with the French data protection authority to assess privacy risks in AI models (PANAME project).
<https://www.cnil.fr/fr/projet-paname-participez-aux-tests-dun-outil-daudit-rgpd-des-modeles-dia>

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)			
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.

Criteo has promoted scientific research in key AI topics via its Criteo AI Lab since its creation in 2018. In particular, Criteo has invested in privacy and fairness research via the publication of several research articles in top-tier scientific conferences, the funding of several PhD students in partnership with renowned French scientific institutions, and the creation of a joint research team, FAIRPLAY, with Inria and ENSAE (<https://team.inria.fr/fairplay/>).

Criteo has also recently collaborated with the French data protection authority (CNIL) to beta-test PANAME, to assess privacy risks in AI models.

Any further comments and implementation documentation.

No answer provided

Section 2 – Risk Management and Information Security

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

Plan and design

Collect and process data

Build, integrate and interpret model

Verify and validate

Deploy

Operate and monitor

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

Criteo's AI Product Development and Deployment Guidelines govern the full lifecycle of AI in Criteo products, whether they are developed internally or by a third-party. Every new AI project undergoes a risk analysis review. Development follows secure SDLC practices; models are validated against dedicated datasets and tested in controlled environments before release. In operation, systems are monitored (including drift and train/serve-skew monitoring and security monitoring), with human oversight maintained and the ability to override, isolate, or switch off any non-compliant AI. Third-party

AI tools require a legally reviewed agreement and internal authorization before use with Criteo data. The Guidelines are reviewed and updated regularly and in case of technological and regulatory change, with approval by the Legal team.

Applicable practices in the system lifecycle stages:

- **Plan and design:** Standard kick-off process for each development project ensures architecture validation, vetting any third-party tool used, and risk analysis (privacy, security, ethics...)
- **Collect and process data:** Verify that data can be used legally and contractually; apply data minimization, anonymization/de-identification/pseudonymization, encryption and appropriate access protections; filter data
- **Build, integrate and interpret model:** Use rigorous development and validation ML methodologies and appropriate statistical techniques; regularly update models to reflect changing conditions; follow standard SDLC practices; verify licensing and confidentiality protections for training data, foundation models and model outputs.
- **Verify and validate:** Test and validate the AI product in a controlled and monitored environment before release; test under a variety of conditions and validate the performance of the AI components; use dedicated datasets for models and agents validation and gather user feedback to identify necessary improvements to agents.
- **Operate and monitor:** Monitor for model drift or sudden change in behavior and train/serve skew continuously where technically possible; maintain the ability to fix observed issues and override, modify, isolate or turn off the system when needed; alerting in place with on-call responsables to fix issues; clients and users have channels to report issues which are processed according to our escalation processes.

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).

Training and validation is relevant to the intended purpose, and protected through data minimization, pseudonymization, encryption and access control. Criteo also maintains procedures to delete requested data across the full data and model pipeline. Criteo does not have access to sensitive attributes relevant for the fairness context.

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

- Network isolation from external access
- Locally hosted testing environments
- Restrictions on outbound Internet or API access, including a list of permitted domains or APIs
- Isolation of staging environments (from production / operational environments)
- Controlled security vulnerability testing
- Human-in-the-loop controls for higher-risk test cases
- Access controls, including roles-based permissions
- Audit logging and monitoring

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.

The CISO and his teams

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

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

- Maintain a copyright policy that addresses data sourcing, content provenance, safeguards, and organizational roles and responsibilities
- Apply provenance controls to guard against training data being sourced from websites or repositories that have been determined to engage in large-scale piracy
- Refrain from the unlawful circumvention of technological protection measures in data sourcing
- Maintain policies or practices to respect copyright holders' preferences through, where appropriate/applicable, mechanisms like Robots.txt "do not train" tags or other machine-readable protocols developed through voluntary, industry-led standards processes
- Conduct data cleaning, qualification, and/or tokenization activities as part of a transformative data preparation process
- Consult and/or retain an intellectual property attorney to protect, enforce, and comply with, intellectual property rights.

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

- Not applicable

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

- Make information about data privacy, practices, and controls accessible to a diverse set of stakeholders
- Use security controls, including encryption at rest and in transit
- Apply Privacy-Enhancing Technologies (PETs) where applicable (e.g. differential privacy, federated learning, synthetic data generation, homomorphic encryption)
- Other(s) (specify): Data pseudonymisation and privacy audits on model weights
- Provide organizational customers with controls to configure and implement privacy choices (e.g., data retention, logging, access, or processing settings)

Describe key privacy measures and data-protection standards applied.

Criteo's Global Data Privacy Policy and privacy governance framework apply group-wide, covering accountability, privacy by design and default, risk management, data subject rights and third-party management. Products, features or changes involving personal data go through a Privacy Compliance Assessment, escalating to a Privacy Impact Assessment (Article 35 GDPR) and a Transfer Impact Assessment where applicable; the DPO reviews outcomes and the security team assess technical controls. Measures include data minimization, pseudonymization, defined retention, regional data residency, role-based access control, encryption in transit and at rest, and as needed exclusion of data from training data and user input fields. Externally shared datasets are anonymized, and differential

privacy is applied so that models cannot disclose training data; model anonymity was tested with the French data protection authority under the PANAME pilot testing. Controls are tested annually under the Privacy Control Plan, mapped to GDPR, ISO/IEC 27701, SOC 2 and NIST Privacy Framework criteria. Independent assurance covers the security control environment.

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

- These responsibilities are delegated contractually to model or infrastructure providers

11.B. Technical and operational controls

- These responsibilities are delegated contractually to model or infrastructure providers

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

- These responsibilities are delegated contractually to model or infrastructure providers

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

Not applicable

12. What is your organization's vulnerability management process? Does your organization take actions to address identified risks and vulnerabilities, including in collaboration with other stakeholders?

- Internal vulnerability disclosure receipt and tracking processes
- Bug bounty or incentive programs for external reporters
- Investigation of reported or identified risks and vulnerabilities
- Development of patches or other mitigations for confirmed risks or vulnerabilities
- Testing of patches or other mitigations
- Scanning for and mitigation of similar risks or vulnerabilities
- Defined escalation criteria and follow-on actions for identified vulnerabilities (e.g., for external reporting or information sharing)

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

- Other(s) (please specify): Not applicable

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.

- Constraints on the agent's action space (e.g., tool-use permissions, environment sandboxing, allowlists/blocklists of resources, scope limits per task)
- Monitoring (anomaly detection, multi-agent interaction monitoring)
- Logging of agent actions

Please provide any further comments and implementation documentation.

No answer provided

Section 3 – Transparency Reporting on Advanced AI Systems

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

Not applicable

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?

- Not applicable

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.

Not applicable

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

Not applicable

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?

Not applicable

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?

Not applicable

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

Not applicable

Section 4 – Organizational Governance, Incident Management, and Transparency

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

- Senior management oversight
- 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).

The product department makes the decision. Final accountability for AI risk-related decisions lies with the leadership team. Policies are reviewed regularly and on demand depending on circumstances

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).

Regular training modules

Channels to report concerns

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.

Reported incidents are documented and maintained according to Criteo's incident management process. Records typically include incident's details (e.g., date & time, affected services or assets, owner, resolution actions, closure status). Lessons learned and management escalation may also be documented and tracked to completion. Access to incident records is restricted to authorized personnel on a need-to-know basis.

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

Yes

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

Most of our research output is made publicly available via scientific publications and presented at top-tier scientific conferences such as NeurIPS, ICML, and ICLR. <https://ailab.criteo.com> Our research work follows open science principles, with freely accessible preprints of our articles and open-source code. Beyond pure research, our engineers and researchers regularly present at external events, write blog posts, and organize conferences where our best practices are shared with institutional and private actors (<https://ailab.criteo.com/trustworthy-ai-symposium>).

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)
- Industry-led best practices or codes of conduct

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

ISO/IEC 27001

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?

- Other (specify): Not applicable

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).

No answer provided

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

Not applicable

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?

- Data quality and representativeness
- System performance and outcomes (e.g. accuracy, consistency)
- Robustness or reliability

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

The above areas have been key research topics within the Criteo AI Lab, including the work of several permanent researchers, funding of several PhD students and interns, creating a joint research team (FAIRPLAY) with renowned French research institutions ENSAE and INRIA. Criteo has published dozens of papers in these areas of research (<https://ailab.criteo.com/publications/>). Most of our research output is made publicly available via scientific publications and presented at top-tier scientific conferences such as NeurIPS, ICML, and ICLR. Our research work follows open science principles, with freely accessible preprints of our articles and open-source code

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

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

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

Criteo provides comprehensive onboarding, regular training, and punctual in-depth AI courses to its employees. Criteo has invested in reducing the environmental footprint of its data centers, e.g. by adopting liquid cooling. More information can be found in our CSR report (https://filecache.investorroom.com/mr5ir_criteo/2576/Criteo-CSR-Report-2025.pdf).

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?

Not applicable

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

Yes

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

Criteo contributes to AI research aimed at societal benefits, particularly fairness and privacy. It has published FairJob, a public dataset supporting research into fairness and bias mitigation in job advertising, and CriteoPrivateAds, a public dataset enabling researchers to develop and evaluate privacy-preserving advertising systems. These resources make real-world benchmarks available to the wider research community, supporting progress toward more equitable access to employment opportunities and better protection of user privacy. Criteo also collaborates with public research institutions, including INRIA, on responsible AI.

Any further comments and implementation documentation.

No answer provided

Generated from the OECD AI Transparency Reporting Framework