Corporate AI Governance Benchmark

A primary-source benchmark of published governance, safety and transparency practices among major AI developers.

Published: Aug 9, 2026 · Evidence date: Aug 9, 2026

Publisher: AI Governance World LLC · aigovernanceworld.com

Executive summary

The launch benchmark covers six major AI developers and technology companies represented in the verified corpus. It inventories official governance frameworks, safety policies, transparency reporting, model documentation and stated regulatory positions under a common evidence window. The charts show documentary coverage only; they do not rank corporate performance.
The public comparability rubric defines six dimensions and four evidence states. Company-authored material is recorded as attributable Tier 3 disclosure, while regulator findings, independent evaluations and incidents remain separate evidence classes. No company receives a score until the evidence for every eligible dimension has completed human review under the same cutoff and the scored edition is explicitly published.

Methodology

The benchmark cohort is the set of verified company entities included in the edition. Eligible evidence consists of official company reports and technical documentation, Tier 1 regulator or legal records and Tier 2 independent evaluations with attributable methods. Media is excluded as primary evidence. All companies use the same evidence cutoff, dimension definitions and source-eligibility rules.
The six comparability dimensions are accountability and oversight; risk governance; model and system documentation; testing and evaluation; incident response and external reporting; and transparency about policy implementation. Each dimension receives an evidence state of documented, partially documented, not documented in the evidence window or not assessed. These states describe evidence availability and are not numeric scores. Any future score requires reviewer approval, source-level citations and a separately published scoring method.

Published governance and risk processes

The benchmark records named frameworks, accountable structures, escalation thresholds, review mechanisms and stated risk domains when the primary source discloses them. General commitments are kept separate from operational procedures and from evidence of implementation.
Legal obligations and regulator findings are linked as external records rather than rewritten as company claims. This preserves both attribution and the distinction between voluntary governance and mandatory compliance.

Transparency and model documentation

System cards, model cards, transparency reports, safety policies and material-update practices are tracked as distinct evidence types. A company-level framework does not automatically populate a model-level evaluation, and a single model card does not establish company-wide practice.
The corpus chart counts verified primary evidence records by company and source type. It does not count words, pages or marketing claims as additional evidence.

Interpretation limits

Public disclosure varies by company, product and jurisdiction. Corporate structures also differ, so a subsidiary model developer and a diversified public company may publish evidence at different organizational levels. Comparability depends on stable definitions, equivalent evidence windows and explicit entity boundaries. The benchmark therefore supports source-level due diligence and gap identification. It is not certification,

investment advice or a finding of legal compliance.

Cohort and evidence window

The initial cohort contains the six verified company records in the launch corpus. Inclusion indicates that a company has an attributable primary source and an entity record; it does not indicate endorsement, superior practice or complete disclosure.
Every comparison uses the edition cutoff and the latest eligible source collected on or before that date. Later publications enter a later update or a documented revision rather than changing the evidence window silently.

Public comparability rubric

The rubric applies six dimensions: (1) accountability and oversight, including named governance structures and decision rights; (2) risk governance, including defined processes for identifying, escalating and treating AI risk; (3) model and system documentation, including intended use, limitations and material changes; (4) testing and evaluation, including documented methods and responsible evaluators; (5) incident response and external reporting, including escalation, notification and corrective-action processes; and (6) implementation transparency, including evidence that published policy is operationalized.
The dimensions are applied consistently to each company. A source can support more than one dimension only when the cited evidence is specific to each claim.

Evidence states, not inferred performance

Documented means eligible evidence directly supports the rubric item. Partially documented means the source addresses only part of the item. Not documented in the evidence window means reviewers found no eligible source supporting the item by the cutoff. Not assessed means evidence collection or review is incomplete.
Not documented is not proof that a practice is absent, and documented is not proof that a practice is effective. The states make disclosure comparability visible without turning public-relations volume into a performance score.

Human review and score publication gate

A numeric benchmark result can be created only after eligible evidence has been archived, mapped to the rubric, reviewed for attribution and assessed under the same cutoff. Reviewer notes must distinguish first-party statements, independent evaluation and enforcement evidence.
This edition publishes the rubric and evidence coverage but no company scores. A later scored edition must publish its weighting, treatment of missing evidence, review protocol and revision policy before rankings appear.

Charts and data

Verified primary evidence records by company

Counts represent attributable primary records, not governance performance.

Anthropic: 1
Google DeepMind: 1
Meta: 1

Microsoft: 1
NVIDIA: 1
OpenAI: 1
Companies in the benchmark cohort grouped by count of verified primary evidence records.

Company evidence by source type

Source form is preserved so reports and model cards are not treated as equivalent evidence.

Company report: 5
Model card: 1
Verified company evidence records grouped by source type.

Structured company evidence coverage

Populated fields show documentary coverage only and do not constitute rubric scores.

Companies in scope: 6
Governance fields populated: 6
Safety fields populated: 6
Transparency fields populated: 6
6 companies are in scope; 6 have a governance field, 6 a safety field and 6 a transparency field populated from verified sources.

References

1. Regulation (EU) 2024/1689 — Artificial Intelligence Act. Source: European Union. Jul 12, 2024. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=OJ:L_202401689

2. SB-53 Artificial intelligence models: large developers. Source: California Legislative Information. Sep 29, 2025. https://leginfo.legislature.ca.gov/faces/billNavClient.xhtml?bill_id=202520260SB53

3. NHTSA Announces Consent Order with Cruise After Company Failed to Fully Report Crash Involving Pedestrian. Source: National Highway Traffic Safety Administration. Sep 30, 2024. https://www.nhtsa.gov/press-releases/consent-order-cruise-crash-reporting

4. OpenAI Frontier Governance Framework. Source: OpenAI. May 28, 2026. https://openai.com/index/openai-frontier-governance-framework/

5. Anthropic Responsible Scaling Policy. Source: Anthropic. Apr 29, 2026. https://www.anthropic.com/rsp-updates

6. Gemini 3.5 Flash Model Card. Source: Google DeepMind. May 19, 2026. https://deepmind.google/models/model-cards/gemini-3-5-flash/

7. 2025 Responsible AI Transparency Report. Source: Microsoft. Jun 20, 2025. https://www.microsoft.com/en-us/corporate-responsibility/responsible-ai-transparency-report/

8. Responsible AI principles at Meta. Source: Meta. Aug 9, 2026. https://ai.meta.com/about/

9. NVIDIA Trustworthy AI. Source: NVIDIA. Aug 9, 2026. https://www.nvidia.com/en-us/ai-trust-center/trustworthy-ai/