

Referință comparativă pentru guvernanța corporativă a IA

O comparație bazată pe surse primare a practicilor publicate de guvernanță, siguranță și transparență ale marilor dezvoltatori de IA.

Publicat: 9 aug. 2026 · Data dovezilor: 9 aug. 2026

Editor: AI Governance World LLC · aigovernanceworld.com

Rezumat executiv

Referința comparativă inițială acoperă șase mari companii de tehnologie și dezvoltatori de IA reprezentați în corpusul verificat. Aceasta inventariază cadrele oficiale de guvernare, politicile de siguranță, raportările de transparență, documentația modelelor și pozițiile declarate privind reglementarea, într-o fereastră comună de dovezi. Graficele prezintă exclusiv acoperirea documentară; ele nu clasifică performanța companiilor. Grila publică de comparabilitate definește șase dimensiuni și patru stări ale dovezilor. Materialele redactate de companii sunt înregistrate ca informări atribuibile de Nivel 3, în timp ce constatările autorităților de reglementare, evaluările independente și incidentele rămân clase distincte de dovezi. Nicio companie nu primește un scor până când dovezile pentru fiecare dimensiune eligibilă nu finalizează analiza umană în cadrul aceleiași date-limită și până când ediția cu scoruri nu este publicată explicit.

Metodologie

Cohorta referinței comparative este formată din entitățile de tip companie verificate și incluse în ediție. Dovezile eligibile constau în rapoarte oficiale și documentație tehnică ale companiilor, evidențe juridice sau ale autorităților de reglementare de Nivel 1 și evaluări independente de Nivel 2 cu metode atribuibile. Presa este exclusă ca dovadă primară. Toate companiile folosesc aceeași dată-limită pentru dovezi, aceleași definiții ale dimensiunilor și aceleași reguli de eligibilitate a surselor. Cele șase dimensiuni de comparabilitate sunt responsabilitatea și supravegherea; guvernarea riscurilor; documentația modelelor și sistemelor; testarea și evaluarea; răspunsul la incidente și raportarea externă; precum și transparența implementării politicilor. Fiecare dimensiune primește una dintre stările: documentată, documentată parțial, nedocumentată în fereastra de dovezi sau neevaluată. Aceste stări descriu disponibilitatea dovezilor și nu sunt scoruri numerice. Orice scor viitor necesită aprobarea evaluatorului, citări la nivelul sursei și o metodă de calcul publicată separat.

Procese publicate de guvernare și gestionare a riscurilor

Referința comparativă înregistrează cadrele nominalizate, structurile responsabile, pragurile de escaladare, mecanismele de analiză și domeniile de risc declarate atunci când sursa primară le publică. Angajamentele generale sunt păstrate separat de procedurile operaționale și de dovezile implementării. Obligațiile juridice și constatările autorităților de reglementare sunt asociate ca evidențe externe, nu reformulate ca afirmații ale companiei. Astfel sunt păstrate atât atribuirea, cât și diferența dintre guvernarea voluntară și conformitatea obligatorie.

Transparență și documentația modelelor

Fișele de sistem, fișele de modele, rapoartele de transparență, politicile de siguranță și practicile privind actualizările semnificative sunt urmărite ca tipuri distincte de dovezi. Un cadru la nivel de companie nu completează automat o evaluare la nivel de model, iar o singură fișă de model nu demonstrează o practică la nivelul întregii companii. Graficul corpusului numără evidențele primare verificate după companie și tip de sursă. Nu transformă numărul de cuvinte, pagini sau afirmații de marketing în dovezi suplimentare.

Limite de interpretare

Informarea publică variază în funcție de companie, produs și jurisdicție. Și structurile corporative diferă, astfel încât un dezvoltator de modele subsidiar și o companie publică diversificată pot publica dovezi la niveluri organizaționale diferite. Comparabilitatea depinde de definiții stabile, ferestre de dovezi echivalente și delimitări explicite ale entităților.

Referința comparativă sprijină, așadar, diligența necesară la nivelul sursei și identificarea lacunelor. Nu reprezintă certificare, consultanță de investiții sau constatare a conformității juridice.

Cohortă și fereastra de dovezi

Cohorta inițială conține cele șase evidențe verificate ale companiilor din corpusul de lansare. Includerea arată că o companie are o sursă primară atribuibilă și o evidență de entitate; nu reprezintă susținere, practică superioară sau informare completă.

Fiecare comparație folosește data-limită a ediției și cea mai recentă sursă eligibilă colectată până la data respectivă. Publicațiile ulterioare intră într-o actualizare viitoare sau într-o revizuire documentată, în loc să modifice în tăcere fereastra de dovezi.

Grila publică de comparabilitate

Grila aplică șase dimensiuni: (1) responsabilitate și supraveghere, inclusiv structurile de guvernare nominalizate și drepturile decizionale; (2) guvernarea riscurilor, inclusiv procesele definite pentru identificarea, escaladarea și tratarea riscurilor IA; (3) documentația modelelor și sistemelor, inclusiv utilizarea prevăzută, limitările și modificările semnificative; (4) testare și evaluare, inclusiv metodele documentate și evaluatorii responsabili; (5) răspuns la incidente și raportare externă, inclusiv procesele de escaladare, notificare și acțiune corectivă; și (6) transparența implementării, inclusiv dovezile că politica publicată este aplicată operațional.

Dimensiunile sunt aplicate consecvent fiecărei companii. O sursă poate susține mai multe dimensiuni numai atunci când dovada citată este specifică fiecărei afirmații.

Stări ale dovezilor, nu performanță dedusă

„Documentată” înseamnă că o dovadă eligibilă susține direct elementul grilei. „Documentată parțial” înseamnă că sursa acoperă doar o parte a elementului. „Nedocumentată în fereastra de dovezi” înseamnă că evaluatorii nu au găsit până la data-limită nicio sursă eligibilă care să susțină elementul. „Neevaluată” înseamnă că procesul de colectare sau analiză a dovezilor este incomplet.

„Nedocumentată” nu dovedește absența unei practici, iar „documentată” nu dovedește eficacitatea acesteia. Stările fac vizibilă comparabilitatea informărilor fără a transforma volumul relațiilor publice într-un scor de performanță.

Analiză umană și pragul de publicare a scorurilor

Un rezultat numeric poate fi creat numai după arhivarea dovezilor eligibile, asocierea lor cu grila, verificarea atribuirii și evaluarea în cadrul aceleiași date-limită. Notele evaluatorilor trebuie să diferențieze afirmațiile proprii, evaluările independente și dovezile de aplicare.

Această ediție publică grila și acoperirea dovezilor, dar niciun scor al companiilor. O ediție ulterioară cu scoruri trebuie să își publice ponderile, modul de tratare a dovezilor lipsă, protocolul de analiză și politica de revizuire înainte de afișarea clasamentelor.

Grafice și date

Evidențe primare verificate după companie

Valorile reprezintă evidențe primare atribuibile, nu performanța guvernantei.

Anthropic: 1

Google DeepMind: 1

Meta: 1

Microsoft: 1

NVIDIA: 1

OpenAI: 1

Companiile din cohorta referinței comparative, grupate după numărul evidențelor primare verificate.

Dovezi despre companii după tipul sursei

Forma sursei este păstrată, astfel încât rapoartele și fișele de modele să nu fie tratate ca dovezi echivalente.

Raport al companiei: 5

Fișă de model: 1

Evidențe verificate despre companii, grupate după tipul sursei.

Acoperirea structurată a dovezilor despre companii

Câmpurile completate arată doar acoperirea documentară și nu constituie scoruri ale grilei.

Companii incluse: 6

Câmpuri de guvernanță completate: 6

Câmpuri de siguranță completate: 6

Câmpuri de transparență completate: 6

6 companii sunt incluse; 6 au un câmp de guvernanță, 6 un câmp de siguranță și 6 un câmp de transparență completat din surse verificate.

Referințe

1. Regulamentul (UE) 2024/1689 — Regulamentul privind inteligența artificială. Sursă: Uniunea Europeană. 12 iul. 2024. Titlul original al sursei: Regulation (EU) 2024/1689 — Artificial Intelligence Act.

https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=OJ:L_202401689

2. SB-53 — Modele de inteligență artificială: dezvoltatori mari. Sursă: Informații legislative ale statului California. 29 sept. 2025. Titlul original al sursei: SB-53 Artificial intelligence models: large developers.

https://leginfo.legislature.ca.gov/faces/billNavClient.xhtml?bill_id=202520260SB53

3. NHTSA anunță un ordin prin consimțământ cu Cruise după raportarea incompletă a unui accident în care a fost implicat un pieton. Sursă: Administrația Națională pentru Siguranța Traficului Rutier. 30 sept. 2024. Titlul original al sursei: NHTSA Announces Consent Order with Cruise After Company Failed to Fully Report Crash Involving Pedestrian. <https://www.nhtsa.gov/press-releases/consent-order-cruise-crash-reporting>

4. Cadrul OpenAI de guvernanță a modelelor de frontieră. Sursă: OpenAI. 28 mai 2026. Titlul original al sursei: OpenAI Frontier Governance Framework. <https://openai.com/index/openai-frontier-governance-framework/>

5. Politica Anthropic de scalare responsabilă. Sursă: Anthropic. 29 apr. 2026. Titlul original al sursei: Anthropic

Responsible Scaling Policy. <https://www.anthropic.com/rsp-updates>

6. Fișa modelului Gemini 3.5 Flash. Sursă: Google DeepMind. 19 mai 2026. Titlul original al sursei: Gemini 3.5 Flash Model Card. <https://deepmind.google/models/model-cards/gemini-3-5-flash/>

7. Raportul de transparență privind IA responsabilă 2025. Sursă: Microsoft. 20 iun. 2025. Titlul original al sursei: 2025 Responsible AI Transparency Report. <https://www.microsoft.com/en-us/corporate-responsibility/responsible-ai-transparency-report/>

8. Principiile IA responsabile la Meta. Sursă: Meta. 9 aug. 2026. Titlul original al sursei: Responsible AI principles at Meta. <https://ai.meta.com/about/>

9. IA de încredere la NVIDIA. Sursă: NVIDIA. 9 aug. 2026. Titlul original al sursei: NVIDIA Trustworthy AI. <https://www.nvidia.com/en-us/ai-trust-center/trustworthy-ai/>