The AI Risk Taxonomy

Abbreviations and Acronyms

All abbreviations and acronyms used in this document are spelt out below, so no prior knowledge is assumed.

AI : Artificial Intelligence

EU : European Union

EU AI Act : European Union Artificial Intelligence Act

A free reference map from AI GOVERNANCE CHAIN™

You cannot govern what you cannot name. Most organizations fail at AI governance not because they lack effort, but because they have no shared map of what can actually go wrong. This taxonomy is that map: thirty named AI risks, organized into six mutually exclusive and collectively exhaustive categories. Print it, share it, and use it to ask the only question that matters at the start of any governance effort: which of these are true for us, right now?

How to use this map

Read each category. For every risk, ask three questions: Could this happen to us? Would we know if it did? Do we have an owner for it? Any risk where the answer is no, maybe, and no is a priority. The free AI Risk Quick-Screen turns this into a fast scored triage.

Category 1: Technical Risks

The system itself fails.

1. Model error and inaccuracy , the system produces wrong outputs within its intended use.
2. Model drift , performance degrades silently as the world changes around a static model.
3. Robustness failure , the system breaks on inputs outside its training distribution.
4. Adversarial manipulation , a bad actor crafts inputs to force a harmful output.
5. Hallucination and fabrication , the system invents facts, citations, or confident falsehoods.

Category 2: Data Risks

The inputs are flawed or mishandled.

6. Training data bias , skewed data teaches the model a skewed view of the world.
7. Data quality failure , incomplete, stale, or mislabeled data corrupts outputs.
8. Privacy violation , personal data is used or exposed without basis or consent.
9. Data provenance failure , the origin and rights of training data cannot be established.
10. Data security breach , training data or prompts are exfiltrated.

Category 3: Ethical and Societal Risks

The system harms people or the public good.

11. Discrimination and unfair outcomes , the system treats groups inequitably.

12. Sycophancy and truthfulness failure , the system agrees, flatters, or tells people what they want to hear instead of what is true.

13. Manipulation and dark patterns , the system nudges people against their interest.

14. Erosion of human autonomy , people defer to the system and stop deciding.

15. Societal and labor disruption , automation displaces work faster than institutions adapt.

Category 4: Legal and Regulatory Risks

The organization breaches its obligations.

16. Regulatory non-compliance , the system violates the EU AI Act, sector rules, or local law.

17. Liability exposure , harm caused by the system creates legal claims.

18. Intellectual property infringement , training or outputs breach copyright or licensing.

19. Contractual and disclosure failure , the organization misrepresents what its AI does.

20. Cross-border conflict , the system meets contradictory rules across jurisdictions.

Category 5: Operational Risks

The organization cannot run or control the system.

21. Lack of human oversight , no one can meaningfully intervene.

22. No off switch , the system cannot be stopped quickly and safely.

23. Vendor and supply-chain dependency , a third-party model or service fails or changes.

24. Monitoring blind spots , failures occur with no detection.

25. Skills and capacity gap , the people running the system cannot govern it.

Category 6: Strategic and Existential Risks

The system threatens the mission or beyond.

26. Reputational damage , a public failure destroys trust faster than any control can rebuild it.

27. Mission misalignment , the system optimizes for a goal that conflicts with the organization's purpose.

28. Agentic over-reach , an autonomous system takes consequential actions beyond intended scope.

29. Concentration and lock-in , dependence on a single model or provider becomes irreversible.

30. Frontier and emergent risk , the most advanced systems behave in ways no one anticipated.

Why this map matters

A taxonomy is not academic. It is the difference between governing by anecdote and governing by coverage. When a board asks "are we exposed?", a named, complete map lets you answer with precision

rather than reassurance. It also organizes everything else in this channel: every framework, playbook, and product maps back to one or more of these thirty risks.

Where the free version stops

This map names the risks and helps you spot them. It does not score your specific exposure, plan treatment, set your risk appetite, or build the monitoring that catches these risks early. That operational depth is the AI Risk Management Kit.

> **Your next step:** Get the **AI Risk Management Kit, Full Edition** in the Execution Library (www.aigovernancechain.com/storefront.html, section: The Core Governance Kits). Every instrument in the Execution Library is also included in the **Governance Vault**, the all-access annual licence.

AI GOVERNANCE CHAIN™

Member, Harvard Business Review Advisory Council.

Document standard: human-first AI governance.