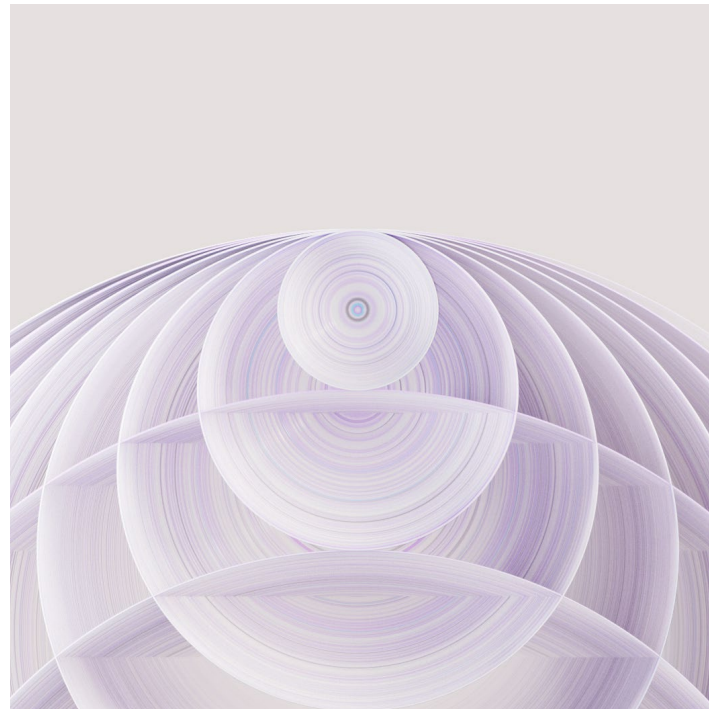


# Manage AI risk

Proactively detect and mitigate risks, monitoring for fairness, bias and drift, and for new generative LLM metrics



Organizations basing decisions on unfair or biased results can face audits, fines, costly litigation and reputational damage. Strategically implementing AI governance is key in monitoring and managing new risks associated with Generative AI (Gen AI) and for machine learning (ML) models. The key to trusted AI includes:

**Monitor for new Gen AI risks.** Gen AI models drive new risk profiles and complexities, and governance can minimize security breaches, privacy/copyright violations, proliferation of misinformation, and misinterpreted context. Governance for both Gen AI and ML models can mitigate bias, drift, inaccurate and unfair results.

**Proactive detection/mitigation of risk.** Organizations that use pre-set alerts to detect when metrics are out of tolerance can proactively mitigate risk. Failure to do so can result in inaccurate, unfair or biased outcomes resulting in brand damage, increased regulatory scrutiny and lawsuits.

**Drive model transparency and documentation.** Transparency improves the ability to track and report on risk at scale. Automating with transparent processes and the capture of model metadata provides the needed documentation for industry and regulatory audits.

IBM® watsonx.governance™ allows you to:

- Monitor and receive alerts for gen AI risks when thresholds are breached for evaluation metrics, instances of toxic language, and use of personally identifiable information (PII) for model inputs and outputs. Monitor for fairness, drift and bias for both gen AI and ML models.
- Automate risk management to identify, measure, monitor and analyze operational risk within a single environment. Integrate risk data, risk and control assessments, internal and external loss events, key risk indicators, and issue and action plan management.
- Gain insights into the state of risk across the enterprise with access to dynamic dashboards and charts used for root cause analysis and more. Drill down through sub-reports for root cause analysis and access browser-based, ad hoc, drag-and-drop report design.

Discover the AI governance solution built for business, with everything you need to identify, manage, monitor and report on risk at scale.

[Learn more](#) →

[Try it for free](#) →

© Copyright IBM Corporation 2024. IBM, the IBM logo, and watsonx.governance are trademarks or registered trademarks of IBM Corp., in the U.S. and/or other countries.

This document is current as of the initial date of publication and may be changed by IBM at any time. Not all offerings are available in every country in which IBM operates.