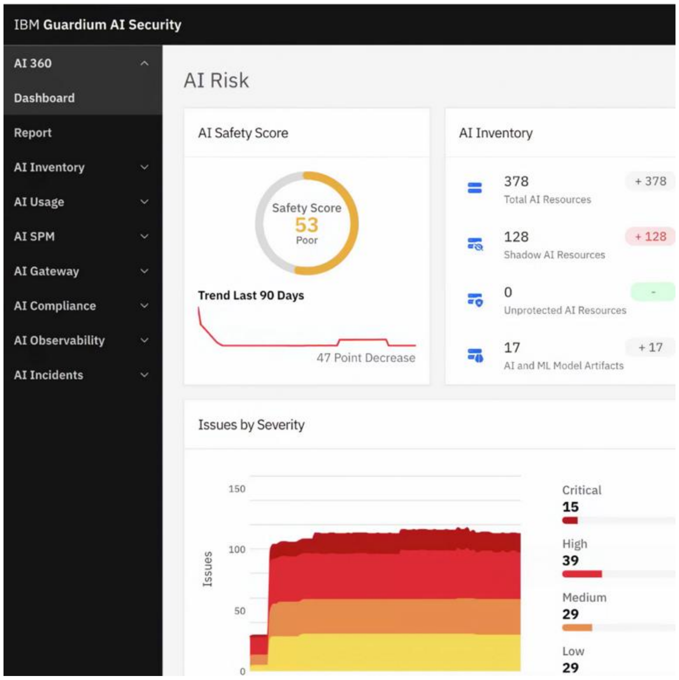
IBM Guardium AI Security

Unlock Trustworthy AI with Guardium AI Security

Secure AI models and AI agents.
Automatically discover shadow AI.
Unify teams for trustworthy AI.



AI Risk

AI Safety Score

Safety Score 53 Poor

Trend Last 90 Days

47 Point Decrease

AI Inventory

378	Total AI Resources	+ 378
128	Shadow AI Resources	+ 128
0	Unprotected AI Resources	-
17	AI and ML Model Artifacts	+ 17

Issues by Severity

Critical	15
High	39
Medium	29
Low	29

According to a study[1], 82% of executives say secure and trustworthy AI is essential to the success of their business, yet only 24% of current generative AI projects are being secured. This leaves a staggering gap in securing known generative AI projects, and when shadow AI and AI agents are added to the mix, the gap is amplified. As your organization races to implement generative AI projects, the security risks need to be better managed.

IBM Guardium® AI Security allows you to discover shadow AI, secure all AI models and use cases, get real-time protection from malicious prompts, and align teams on common set of metrics—for secure and trustworthy AI.

With IBM Guardium AI Security, you can:

- Discover shadow AI to get full visibility into AI use cases to discover generative AI, shadow AI and agentic AI. Enable automated and continuous monitoring for AI models in your cloud, code repositories and embedded AI.
- Detect security vulnerabilities and misconfigurations and run automated penetration tests. Easily map the results to common assessment frameworks such as OWASP Top 10 for LLM, NIST AI RMF for easy remediation.
- Monitor prompts to your AI and define the security policies. This allows you to better secure your AI application.
- Manage compliance with Guardium AI Security and watsonx.governance. It allows you to look at the same AI inventory, have both security risks and compliance considerations on a single dashboard, and shadow AI insights—for implementing a safe and trustworthy AI.

Guardium AI Security offers a robust, enterprise-grade solution to manage the security of your AI assets, including AI agents, and bring together security and governance teams on a single set of metrics, for secure and trustworthy AI.

It allows you to secure the models you build locally, use on the cloud or consume. With the AI firewall, you can scan and protect the different prompts that are coming into your applications, and the output the AI application is generating. Since each organization is unique, you can easily set the policies for different actions to be triggered for actions like code injection, PII exposure, data leakage, and more.

An out-of-the-box integration with IBM watsonx.governance provides a true Risk and Governance solution for disparate teams to look at a single set of metrics for business and security risks. Now your Governance and Security teams look at the same AI inventory for Trustworthy AI. It offers industry's first software to bring AI security and AI governance teams together and provide a unified view of enterprises' risk posture.

Sign up for a deep-dive demo of Guardium AI Security by registering here, or by scanning this QR code.

[1] Securing Generative AI , May 2024

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.