

Test, evaluate, and certify AI models for trust

Only 23% of IT leaders are very confident in their organization's ability to manage security and governance when rolling out GenAI tools.
Source: Gartner survey of 360 IT leaders (May–June 2025), reported Oct 6, 2025.

As AI moves into high-stakes operations and decisions, government agencies and large enterprises often lack clear visibility into bias, reliability, provenance, and adversarial behavior. Leaders are accountable for outcomes, yet many do not have a consistent way to measure performance and risk across the AI lifecycle, making deployments harder to validate and trust in practice.

Three core challenges

Choosing the right model for each mission

Model selection is often driven by vendor claims and public benchmarks instead of rigorous tests on your own data and use cases.

Knowing whether tuning made things better or worse

Fine-tuning and prompt changes alter behavior, but teams rarely know exactly where performance improved, where it regressed, or how risk shifted.

Proving ongoing reliability, compliance, and governance

Blind spots in drift, bias, and adversarial behavior leave leaders exposed without an audit-ready trail for boards and regulators and governance teams.

Why the status quo is not enough

⚠ OPERATIONAL BLIND SPOTS

Context-free, untested models mislead users and jeopardize integrity and security.

⚠ IDEOLOGICAL BIAS

Most models embed political and cultural biases which can distort truth and sow division.

⚠ INADEQUATE BENCHMARKS

Generic benchmarks and leaderboards miss business-critical requirements, making AI harder to validate and trust.



Solving the AI trust problem across industries



GOVERNMENT | DEFENSE | FINANCE | TELECOMS | SUPPLY CHAIN

From generic benchmarks to use-case-ready results

SeekrGuard is an evaluation and governance capability that tests, monitors, and documents model performance and risk using your organization's context as the benchmark, enabling audit-ready decisions across selection, tuning, and ongoing operations.

SeekrGuard core capabilities

✔ Model Scorecards

Rank options across tasks, datasets, and scenarios for faster model selection.

✔ User-Defined Risk Profiles

Weigh what matters most by mission or business line and standardize governance criteria.

✔ Configurable Evaluators

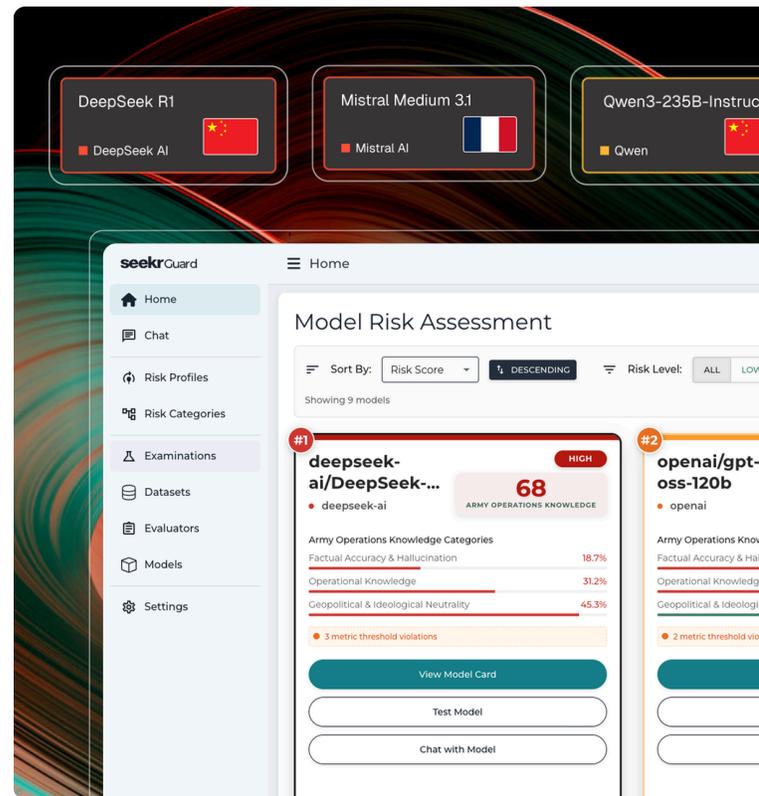
Test edge cases, safety constraints, and adversarial prompts with consistent scoring.

✔ Automatic Test Data Generation

Turn internal documents and knowledge sources into test criteria and datasets, without heavy upfront work.

✔ Flexible Testing

Evaluate open-weight and proprietary models in approved environments.



The Seekr Difference

Built for environments where failure isn't an option

Seekr is a leader in explainable and trustworthy artificial intelligence designed to power mission-critical decisions in enterprises, government, and regulated industries. We provide secure, auditable AI solutions tailored to sectors where transparency, accuracy, and compliance are paramount.

**SOC 2 TYPE II COMPLIANT | FLEXIBLE, SOVEREIGN, SECURE DEPLOYMENT |
3X MORE ACCURATE MODEL RESPONSES | 6X GREATER BUSINESS RELEVANCE**