cyberhaven

AI Risk Management Framework Checklist

Five operating pillars and four accountability requirements to assess how ready your organization is to fully adopt and implement agentic AI.

Every AI approval a security team makes might look reasonable on its own, but individually reasonable decisions add up to enterprise-wide exposure fast. Use this checklist to see where your AI risk management program already has coverage and where the gaps sit, pillar by pillar.

Pillar 1

AI Usage and Shadow AI Discovery

Maintain visibility into every AI tool your organization touches, sanctioned or not.

- [ ] We maintain a continuous inventory of sanctioned and unsanctioned AI tools.
- [ ] Our inventory includes embedded copilots and agents that employees adopt on their own.
- [ ] We track usage intensity, not just adoption rate, across AI tools.
- [ ] We flag smaller, fast-growing tools, such as endpoint agents, that may carry an outsized share of data exposure.

Pillar 2

Sensitive Data and Lineage

Know where your data originated, how it moved, and where it shaped an AI-generated output.

- [ ] We can trace where sensitive data originated before it reached an AI tool.
- [ ] We can trace how data moved across systems, prompts, and sessions.
- [ ] We can identify where specific data influenced an AI-generated output.
- [ ] We assign ownership for derivative content, not just for the original source documents.

Pillar 3

AI-Aware Security Policies

Replace static allow-or-block rules with policies calibrated to risk.

- [ ] Our AI policies are risk-based rather than binary allow-or-block rules.
- [ ] Policies account for data sensitivity.
- [ ] Policies account for tool type.
- [ ] Policies account for user role.
- [ ] Policies account for decision impact.

Pillar 4

Point-of-Use Enforcement

Apply controls while a user interacts with AI, not after data has already left the organization.

- ☐ We enforce safeguards at the moment of AI interaction, not after the fact.
- ☐ Controls apply consistently across endpoints, browsers, and cloud applications.

Pillar 5

Continuous Monitoring and Improvement

Treat AI security as an ongoing cycle. A model that behaves safely today can behave differently tomorrow.

- ☐ We monitor AI systems on an ongoing basis, not only at launch.
- ☐ We investigate anomalies as data, prompts, and adversarial pressure change.
- ☐ Monitoring results feed back into policy and governance updates.

Governance and Accountability Checklist

A framework only holds up when responsibility is explicit.

- ☐ **Risk ownership:** leadership has assigned clear authority to approve, restrict, or terminate AI operation.
- ☐ **Policy and standards:** acceptable use, safety thresholds, and security requirements are defined and tied to legal and operational obligations.
- ☐ **Independent assurance:** reviewers separate from system developers and operators conduct audits and adversarial testing.
- ☐ **Incident response authority:** leadership has defined escalation paths and decision rights for unexpected AI behavior.

Quick NIST Alignment Reference

The NIST (National Institute of Standards and Technology) AI Risk Management Framework organizes AI risk into four functions. Here is how this checklist maps to each one.

NIST Function	What It Covers	Supported By
Govern	Accountability and risk tolerance	AI-aware policies, governance checklist
Map	Where and how AI systems are used	Discovery and lineage visibility
Measure	Risk evaluation through testing and monitoring	Continuous monitoring
Manage	Safeguards and risk tracking over time	Point-of-use enforcement

Want the Full Framework?

This checklist covers the essentials. For the complete five-pillar operating model, mapped against the NIST Cybersecurity Framework (CSF) 2.0 and its AI Profile, see **"Securing AI Systems: A Comprehensive Framework for Enterprise Defense."**