

AI Agent Governance One-Pager




Why Governance Matters



Effective governance ensures your AI agents are safe, ethical, and aligned to business goals—minimizing risk while maximizing impact.

Long-Term Context

- **Governance as Your North Star:** Treat this page as the living playbook for safety, ethics, and reliability—revisit it at every project milestone.
 - **Onboarding New Team Members:** Share this one-pager with any new collaborator to get them up to speed on your standards.
 - **Audit & Compliance Checks:** Use when preparing internal or external reviews—everything you need is here.
-

Key Pillars & Why They Matter



1.  **Data Privacy & Compliance**
 - **What:** Secure PII, follow GDPR/CCPA/HIPAA, enforce retention policies. See NIST's AI Risk Management Framework ([NIST RMF](#))
 - **Why:** Avoid fines, protect customer trust, and pave the way for future data-driven initiatives.
 - **6-Month Check:** Confirm data-handling SOPs are still enforced; update policy for any new regulations.
2.  **Human-in-the-Loop**
 - **What:** Defined review workflows, escalation paths, clear ownership. Based on principles from the Partnership on AI
 - **Why:** Balances speed with safety—people catch what AI can't.
 - **6-Month Check:** Ensure roles haven't drifted and that someone is still actively auditing outputs weekly.
3.  **Bias & Fairness**
 - **What:** Regular bias audits, fairness metrics, mitigation plans. Aligned with OECD's AI Principles (OECD AI)

- **Why:** Keeps your agents equitable and your brand reputationally safe.
 - **6-Month Check:** Run a new bias report on your latest data; adjust thresholds as your use cases evolve.
4.  **Security & Access Control**
- **What:** Vaulted secrets, RBAC, API protection, network segmentation. See ISO/IEC JTC 1/SC 42 guidelines (ISO SC42)
 - **Why:** Prevents breaches and unauthorized actions—critical as adoption scales.
 - **6-Month Check:** Rotate keys, review access logs, and audit user permissions.
5.  **Auditability & Transparency**
- **What:** Comprehensive logs, version control of models/data, periodic reviews. Reflects EU's "Ethics Guidelines for Trustworthy AI" (EU Guidelines)
 - **Why:** Enables root-cause analysis, supports compliance, and builds stakeholder confidence.
 - **6-Month Check:** Archive logs, verify version tags, and present a governance report to leadership.
-

Best Practices for Ongoing Value

- **Embed Early & Often:** Bake privacy, security & fairness checks into your design sprints and retrospectives—don't wait until rollout to think about governance.
- **Iterate & Monitor:** Treat governance as a living process: regularly revisit policies, run fresh bias scans, and update guardrails as your agents evolve.
- **Empower Stakeholders:** Define clear ownership—who approves, who reviews, and who acts on findings—and rotate a "Governance Champion" each quarter to lead reviews and training.
- **Automate Monitoring:** Use tools (e.g., bias scanners, log aggregators) to surface issues before they become crises and feed them back into your sprint backlog.
- **Document & Share:** Maintain a public "Governance Updates" changelog (in your repo or team wiki) so every change is visible and your team stays aligned.

Quick Checklist for Self-Assessment

 / 	Item	6-Month Action
	PII handling plan in place	Audit SOP; update for new data sources
	Human reviewer assigned	Confirm reviewer is active; rotate if needed
	Bias assessment scheduled	Run fresh bias audit; compare to baseline
	RBAC configured	Review permissions and rotate keys
	Logging & versioning enabled	Archive logs; tag new model/data versions

Tip: Store this table in your repo's root README so it's the first thing anyone sees when they land on your project.