

The AI governance maturity model

From ad hoc oversight to adaptive, operational governance

The five levels of AI governance maturity

AI governance can't remain a static policy exercise. Mature organizations operationalize governance across the full AI lifecycle, from model development and evaluation to deployment, monitoring, and auditability.

LEVEL 1

Ad hoc

Reactive, inconsistent, and manual.

WHAT IT LOOKS LIKE

Models and agents deploy without approval. Reviews follow incidents, with no shared definition of safe to ship.

KEY RISKS

- Silent production failures
- No audit trail

NEXT MOVE

Inventory every model in production.
Document one approval process.

LEVEL 2

Documented

Policies exist, but compliance varies.

WHAT IT LOOKS LIKE

Written policies exist with designated reviewers, but evaluation criteria are applied inconsistently across teams.

KEY RISKS

- Policy theater at scale
- Manual review bottlenecks

NEXT MOVE

Automate eval scoring. Embed approvals in deployment tooling.

LEVEL 3

Operational

Governance lives in your workflows.

WHAT IT LOOKS LIKE

Risk tiers drive reviews, evaluation gates precede releases, and model lineage with traces is consistently tracked.

KEY RISKS

- Manual approvals bottleneck growth
- Agents and multi-model systems

NEXT MOVE

Automate evaluation scoring. Embed approvals in deployment tooling.

LEVEL 4

Automated

Evaluation and oversight at scale.

WHAT IT LOOKS LIKE

Automated evaluation pipelines use LLM judges and red-teaming while production monitoring detects model drift.

KEY RISKS

- Automation debt from lagging evals
- False confidence in automated

NEXT MOVE

Feed production signals back into evaluation design. Add human gates for highest-risk decisions.

LEVEL 5

Adaptive

Governance evolves to unlock speed.

WHAT IT LOOKS LIKE

Governance adapts to model and regulatory changes while production signals continuously improve evaluations.

KEY RISKS

- Complexity requires ongoing maturity
- Ongoing investment at scale

NEXT MOVE

Treat governance as infrastructure. Invest in it like compute or reliability.

Key insights

AI governance maturity isn't about adding more checkpoints. It's about making governance observable, repeatable, and operational across the full AI lifecycle.

Production failures occur when evaluations, traces, and reviews are in separate tools. Monitoring ceases at deployment. Approvals exist outside the dev workflow.

Want the full framework?

Read the W&B ebook [Governance workflows for AI agents](#) to help uplevel your team.

Where are you today? Use these questions to unlock your current level.

- ☐ Do you know which AI systems run in production?
- ☐ Can you trace lineage from data to deployment?
- ☐ Are evaluations standardized before every release?
- ☐ Are approvals embedded in CI/CD workflows?
- ☐ Can teams monitor model behavior and drift in prod?
- ☐ Do you comply with NIST AI RMF and the EU AI Act?
- ☐ Can you produce complete audit trails on demand?
- ☐ Does governance help your teams move faster?