Integration Guide

Adapting Runtime Enforcement to Existing Orchestrators

Strategic goal: Reduce engineering resistance

Adoption Philosophy

Runtime enforcement should be:

- incremental
- reversible
- low risk

Most teams start in **observe-only** mode.

Interception Points

The PEP is placed:

- before inference
- before retrieval
- before tool execution
- before state mutation

No changes to business logic required.

Safe Execution Patterns

Two-Phase Execution

1. Model proposes a plan
2. Runtime gates the plan
3. Execution occurs only after verification

Blast-Radius Caps

Hard ceilings on:

- tool calls per session
- affected entities
- write operations

Failure Semantics

Fail-Closed (Writes)

If enforcement is uncertain:

- writes are blocked
- system degrades safely

Fail-Open (Reads)

Non-sensitive reads may proceed with evidence flags.

Migration Path

1. Evidence capture only
2. Policy simulation
3. Selective enforcement
4. Full runtime control

Why Teams Adopt This

Because it:

- prevents incidents instead of explaining them
- simplifies audits
- centralizes responsibility
- removes fear from deployment