AI SECURITY ENGINEERING FIELD GUIDE · 2026

Chapter 08 · AI-Aware Secure SDLC

Standalone learning module for LMS delivery and required reading.

FORMAT	USE	SCOPE	AUDIENCE
Standalone PDF	LMS reading module	Single chapter	Learners

CHAPTER 08

AI-Aware Secure SDLC

INSPECT

Intake, design review, threat-model triggers, eval gates, exceptions, backlog, and retest evidence.

DECISION

A launch gate is useful only when it changes release behavior.

> AI security work fails when it arrives after product decisions are already locked.
>
> FIELD GUIDE

> **FIELD GUIDE: DELIVERY REVIEW**
>
> Inspect how AI features enter intake, design, threat modeling, testing, release gates, exceptions, telemetry, and remediation.

This domain covers secure delivery for AI features: intake, requirements, design review, threat modeling, release gates, eval gates, model/provider changes, prompt changes, retrieval changes, tool changes, feature flags, exceptions, and backlog ownership. It matters when product teams ship AI behavior faster than security, GRC, privacy, and operations can verify it.

The secure SDLC question is whether AI-specific changes trigger the right work before release. Prompt edits, provider swaps, retrieval-source changes, new tools, new memory behavior, and eval failures should be visible as engineering events, not informal product choices.

> **FIELD USE**
>
> Pick a recent AI feature and reconstruct its path from idea to production. The review should show intake, threat model, tests, release gate, exception status, owner, telemetry, and retest plan.

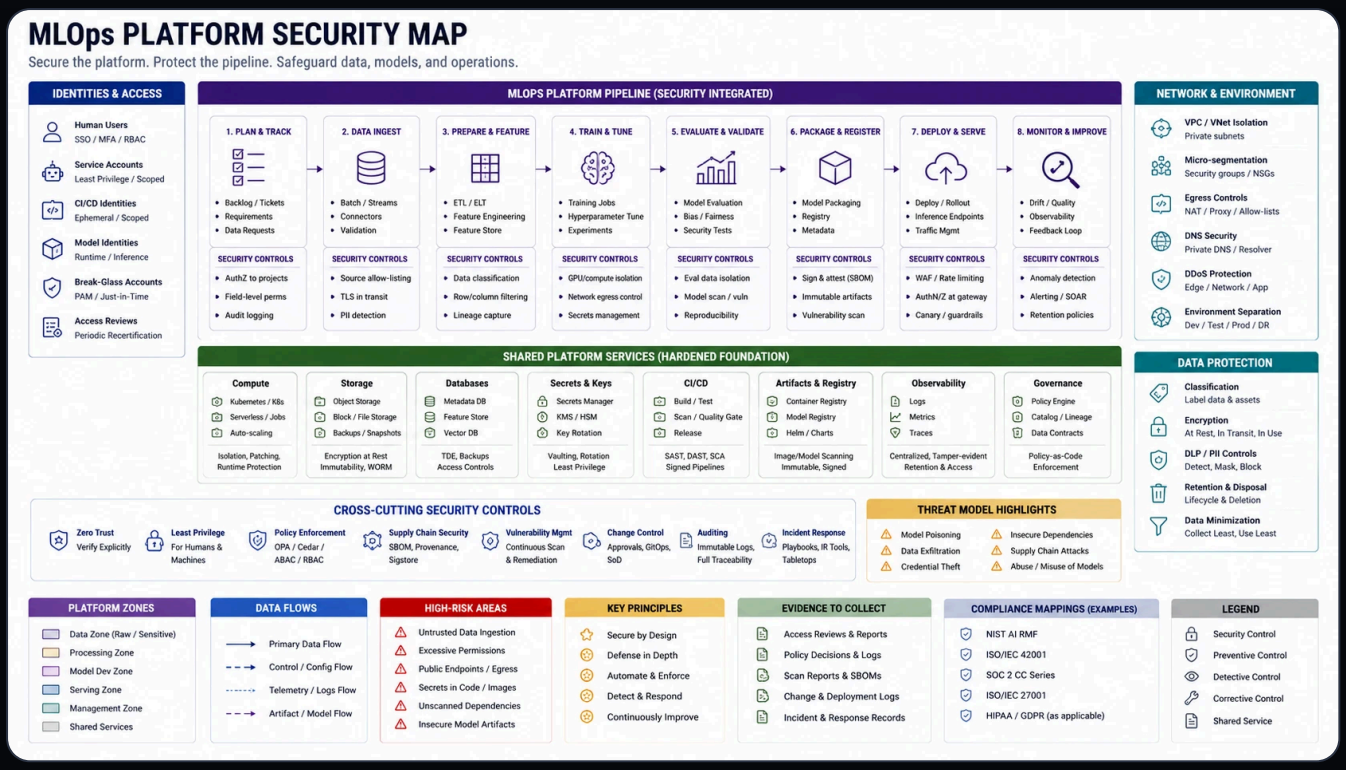

FIGURE 1: FIGURE 9: AI-AWARE SECURE SDLC RELEASE GATE CONNECTING INTAKE, THREAT MODELING, EVAL GATES, EXCEPTION WORKFLOW, AND EVIDENCE REQUIREMENTS

WHAT BREAKS HERE?

- AI features bypass product security review.
- Prompt, model, retrieval, and tool changes ship without security triggers.
- Evals are treated as quality checks but not release gates.
- Exceptions have no owner, expiry, or compensating control.
- Findings do not become backlog work with retest evidence.
- Product requirements omit data class, authority, abuse paths, and evidence needs.

WHAT TO INSPECT

- Product intake forms, design docs, PRDs, and launch checklists.
- Threat model triggers for AI-specific changes.
- CI/CD gates, eval gates, feature flags, and release approvals.
- Exception records, risk acceptances, backlog items, and retest dates.
- Ownership across product, engineering, security, privacy, legal, GRC, and SOC.
- Change records for prompts, models, providers, tools, retrieval sources, and memory.

WHAT TO ASK

- Which AI changes require security review?
- Where do model, prompt, retrieval, tool, and provider changes get recorded?
- What test evidence blocks a release?
- Who owns remediation after a finding?
- How do exceptions expire?
- What telemetry proves a shipped control still works?
- What product requirement would change if the security review happened earlier?

WHAT TO TEST

- Submit a mock AI feature through intake and see whether review triggers fire.
- Change a prompt, model, retrieval source, or tool scope and verify required gates.
- Inspect recent releases for missing AI security evidence.
- Trace one finding into backlog, fix, retest, and closure.
- Review exceptions for owner, expiry, compensating control, and evidence.
- Verify feature flags and rollback paths for high-risk AI behavior.

CONTROLS AND GUARDRAILS

- AI feature intake with risk tier, data class, authority, provider, and evidence fields.
- Threat modeling triggers for AI system changes.
- Release gates for prompt injection, RAG authorization, agent authority, privacy, logging, and eval coverage.
- Exception workflow with owner, expiry, compensating controls, and retest plan.
- Remediation backlog fields for severity, control, evidence, owner, and due date.
- Change-management triggers for model, prompt, provider, retrieval, tool, memory, and telemetry changes.

EVIDENCE TO COLLECT

- AI intake records.
- Threat model artifacts.
- Release gate results.
- Eval and regression test evidence.
- Exception register.
- Backlog and retest records.
- Launch approval notes.
- AI change log and rollback evidence.

OUTPUT ARTIFACTS

- AI-aware SDLC review.
- Release gate checklist.
- AI security backlog.
- Exception register review.
- Control-to-evidence map.

RELATED SERVICES AND WORKBENCH TOOLS

TYPE	RELATED PATHS
Workbench	Threat Canvas, AI Control Crosswalk, Runtime Proxy
Services	Product Security Baseline, AI Product Security Assessment, AI Security Operating Model
Handbook	Threat Modeling, Evaluation and Regression Testing, Governance Evidence and Customer Trust

DECISION - CONDITIONAL: RELEASE GATE DECISION

An AI feature is not ready when the model behaves well once. It is ready when the team can show the review, tests, controls, logs, exception status, and retest path.