B SplendensLabs

✦ CHECKLIST · AI SECURITY

Ship AI that's secure by design.

A pragmatic security checklist for production AI — data isolation, PII handling, prompt-injection defense, access control and audit.

Isolation | PII | Injection | Access | Audit

splendenslabs.com · hello@splendenslabs.com

Before you ship

The production AI security checklist.

Work through every item before exposing an AI system to real users or data. Security is a design property, not a final review.

- ✓ Data isolation — tenants and users cannot reach each other's data.
- ✓ PII — sensitive data is minimized, masked or excluded from prompts.
- ✓ Retrieval — inherits the user's permissions; no over-broad access.
- ✓ Prompt injection — untrusted content is treated as data, not instructions.
- ✓ Tools — every agent tool is least-privilege and validated.
- ✓ Access — role-based permissions on every action and dataset.
- ✓ Secrets — no credentials in prompts, logs or client bundles.
- ✓ Audit — every consequential action is logged and traceable.
- ✓ Kill switch — any agent or flow can be paused instantly.

PRINCIPLE

The goal is not uncontrolled autonomy — it is governed autonomy. Every action explainable, bounded and reversible.