CASE STUDY · SOFTWARE & OBSERVABILITY

Standardised how an engineering org ships AI features — every one instrumented for evaluation and production observability before release

Datadog — Multi-week programme

We delivered an AI development lifecycle programme for a leading observability company's engineering teams — establishing a single governed path from prototype to production in which every AI feature is evaluated and instrumented for observability before release, cutting the path to a production-ready feature by an estimated 40%.

THE CHALLENGE

A high-calibre engineering organisation wanted to accelerate how its teams build and ship AI features — but ad-hoc experimentation meant inconsistent evaluation, patchy observability and no shared bar for production-readiness. For a company whose own product is about measuring systems in production, shipping AI features that could not themselves be measured, evaluated and monitored was untenable.

WHAT WE BUILT

- Ran an AI development lifecycle programme tailored to an experienced engineering audience
- Covered the full path from prototype to production: prompt and agent design, RAG, evaluation harnesses, observability and guardrails
- Emphasised measurability and observability of AI systems in production — a natural fit for the org's domain — so every feature is monitorable once live
- Delivered reference patterns and a lifecycle the teams could standardise on

OUTCOMES

- Cut the path to a production-ready AI feature by an estimated 40%, with every feature now evaluated and instrumented for production observability before release
- Engineering teams aligned on a single repeatable AI development lifecycle
- Evaluation and observability built into how AI features ship rather than added after incidents
- A more confident, governed path from prototype to production

STACK