

GenAI Interview Question Bank

Five questions that separate people who've shipped from people who've watched tutorials — and what strong answers sound like.

1. Design RAG over messy PDFs, scans, and tables.

Strong answers cover: ingestion (OCR, table extraction, layout-aware chunking), hybrid retrieval (dense + keyword + reranking), metadata filtering, evals on a golden set — and where it breaks in production.

2. Production answers are wrong. How do you debug it?

Strong answers cover: trace replay (what was retrieved, what the model saw), separating retrieval failures from generation failures, golden-dataset regression, and the eval harness that should have caught it.

3. Cut inference cost 10× without harming quality.

Strong answers cover: measuring cost per workflow first, smart routing (small models for easy calls), caching, batching, prompt compression — and proving quality held via evals.

4. Let the AI write to your CRM — safely.

Strong answers cover: least-privilege tool permissions, schema validation, idempotency, human-in-the-loop on high-stakes writes, full audit trail, rollback plan.

5. Deliver a two-week pilot for a skeptical customer.

Strong answers cover: ruthlessly narrow scope, agreed success criteria up front, evals from day one, and a production-architecture sketch so the pilot doesn't become throwaway — pilots prove the tip, not the iceberg.