

The 30-Day AI-Engineer Roadmap

The exact path from “I can write Python” to interview-ready on the skills employers actually test — **RAG, agents (LangGraph / MCP), evals, LLMOps, guardrails and system design**. Applied engineering, not research: no heavy math, a portfolio over a certificate. Foundations → Applied → Advanced, every day on the same loop that makes it stick. This page shows you where to start; the next three show exactly what you build, why it matters, and the interview it unlocks.

START WHERE YOU ARE – PICK THE LINE THAT SOUNDS LIKE YOU

“I can already code — but the LLM stack still feels like magic.”

Start at Week 1. The plain-English intuition and your first working RAG make it concrete within days — no math wall.

“I know the theory — but not what actually ships.”

Move fast through the basics and spend your energy where interviews live: retrieval quality, agents, evals and serving.

“I just need the straightest line to interview-ready.”

45–60 minutes a day, one loop — learn, build, rehearse. It’s already sequenced for you; just show up daily.

You’re **not** starting over.

Bring whatever you already have — Python, a little machine-learning theory, maybe a side-project you deployed. This roadmap is the **applied layer** on top of it: the slice of AI engineering that interviews actually test, which most syllabi haven’t caught up to yet. So the newer topics aren’t a gap you’re behind on — they’re the edge you’re building. At every stage we flag what to **lean on** from what you know, what’s **genuinely new**, and what you can **skip** — so it reads as the shortest path, not a second degree.

THE 30 DAYS AT A GLANCE

DAYS 1–7

Foundations & first RAG

DAYS 8–14

Retrieval & reliable calls

DAYS 15–21

Agents & evals

DAYS 22–28

Serving & scale

29–30

Ship & rehearse

► START TONIGHT

DEPLOYED + INTERVIEW-READY ◀

EVERY DAY RUNS ONE LOOP

01 · LEARN

Retrieve, don’t re-watch

100+ active-recall cards, each with an AI tutor and a deep dive; spaced repetition schedules the next hit so it sticks.

02 · BUILD

Write it by hand

39 hands-on labs in an in-browser Python IDE — from the primitives to production services. Code that fails, then fix.

03 · REHEARSE

Say it out loud

AI voice mock interviews plus 5 role-tailored tracks — spoken, scored in real time, calibrated to your target.

WEEK 1 Foundations & your first RAG

Days 1-7

- **Core intuition:** tokens, embeddings, attention — no heavy math
- **Ingest** → **clean** → **enrich** real documents with metadata
- **Chunk & embed**; index in-memory, then FAISS
- **Retrieve** → **prompt** → **cite** grounded answers
- **Refuse** when the corpus doesn't know (no hallucinating)
- **Ship answer()** — your first end-to-end RAG pipeline

WHY NOW

RAG is the one architecture almost every company has actually shipped — so it's the safest, most-asked thing in an AI-engineering interview, and the fastest proof you've built something real.

UNLOCKS

"Walk me through a RAG pipeline end to end — and tell me where it breaks."

SHORTCUT

Already know embeddings & attention from theory? Skim the recap and spend the time on the part courses skip: making retrieval actually work on messy, real documents.

WEEK 2 Retrieval quality & reliable LLM calls

Days 8-14

- **BM25** → **semantic** → **hybrid** (RRF / EnsembleRetriever)
- **Cross-encoder reranking** to fix rank-1
- **Query rewriting** + a faithfulness guard (LCEL)
- **Prompts & roles**, few-shot, output parsers
- **Structured output:** Pydantic + `.with_structured_output()`
- **Tools & function calling:** hand-write the call loop

WHY NOW

"My demo works but the answers are wrong" is where 90% of real systems — and interviews — actually live. Being fluent about *why* retrieval fails is the strongest senior signal there is.

UNLOCKS

"Your answers are wrong — is it a retrieval problem or a generation problem?"

SHORTCUT

Comfortable with typed Python? Pydantic structured output will feel natural — lean on it. New to search? Hybrid + rerank is the single highest-leverage thing you'll learn all month.

WEEK 3 Agents & making it trustworthy

Days 15–21

- **Agents on LangGraph:** a cyclic StateGraph + guardrails
- **RAG-as-a-tool;** multi-tool routing, step caps
- **MCP:** stand up a server + client, progressive tool discovery
- **Memory & conversational RAG** (history-aware rewriting)
- **Eval design:** golden set + the RAG Triad (RAGAS-style)
- **LLM-as-judge,** tracing/observability, a CI eval gate

WHY NOW

Agents and evals are the 2026 hiring frontier — and evals are the most under-taught, most-demanded skill in the field. This is exactly the gap between what most courses cover and what a job needs, so fluency here is an edge, not a deficit.

UNLOCKS

“How do you stop an agent looping forever?” · “Walk me through an eval you designed.”

SHORTCUT

If most of this is new to you, good — it’s new to almost everyone. Nobody expects prior exposure to MCP; showing up fluent is what makes you memorable.

WEEK 4 Guardrails, serving & scale

Days 22–28

- **Guardrails & safety** mapped to the OWASP LLM Top 10
- **PII detection/redaction,** injection & jailbreak defense
- **Serve it:** FastAPI + token streaming (SSE/async)
- **Fast & cheap:** semantic caching, model routing, fallbacks
- **Managed vector DB** (pgvector/Qdrant), ingestion at scale
- **Docker → deploy → on-call;** prompt vs RAG vs fine-tune (LoRA/PEFT)

WHY NOW

This is the “have you actually shipped?” layer — it turns a notebook demo into something an interviewer believes runs in production, and it’s where cost and latency questions get real.

UNLOCKS

“How would you cut latency and cost in production?” · “Prompt, RAG, or fine-tune — when?”

SHORTCUT

Shipped a web service before? FastAPI + Docker will feel familiar — port that instinct straight over. If not, this is the most transferable skill here: it applies to any backend role, not just AI.

DAYS 29–30 Ship the capstone & rehearse for real

The finish

- **Finish the 5-part RAG capstone:** build → agents → evals → deploy
- **Put it in your portfolio** — a deployed project beats a certificate
- **Run AI voice mock interviews** — timed, scored, out loud
- **Unlimited tailored interviews** calibrated to your target role
- **Drill the real-gap question banks** until the answers feel easy
- **Walk in ready** for the questions that end interviews

WHY NOW

A deployed project you can defend plus reps saying it out loud is what converts knowledge into offers. Knowing the material isn't the bottleneck — being able to talk it under pressure is.

UNLOCKS

The whole interview — because by now you've already rehearsed it, out loud, and scored.

What you'll have by day 30

- 1 A deployed RAG app**
Retrieval → agents → evals → FastAPI — a real project you can demo and talk through.
- 2 Primitives built by hand**
From a softmax and attention to hybrid retrieval and a LangGraph agent — not copied.
- 3 Interview reps on the board**
Scored AI voice mocks + unlimited role-tailored drills, out loud and timed.
- 4 Cold-ready answers**
The RAG, agent, eval & system-design questions that end interviews — handled.

INTERVIEW SIGNALS YOU'LL BE ABLE TO ANSWER COLD

- „ Build an agent that decides *when* to retrieve.“
- „ How do you stop an agent looping forever?“
- „ What is MCP, and why does it matter?“
- „ Walk me through an eval you designed.“

Start Week 1 tonight — free.

Foundations is free forever: the real platform, active-recall cards, the built-in IDE and AI grading. No card required. Start now at skillumen.com — prefer a 30-day sprint or a self-paced pace? Either way, the path above is the same.

skillumen.com →