

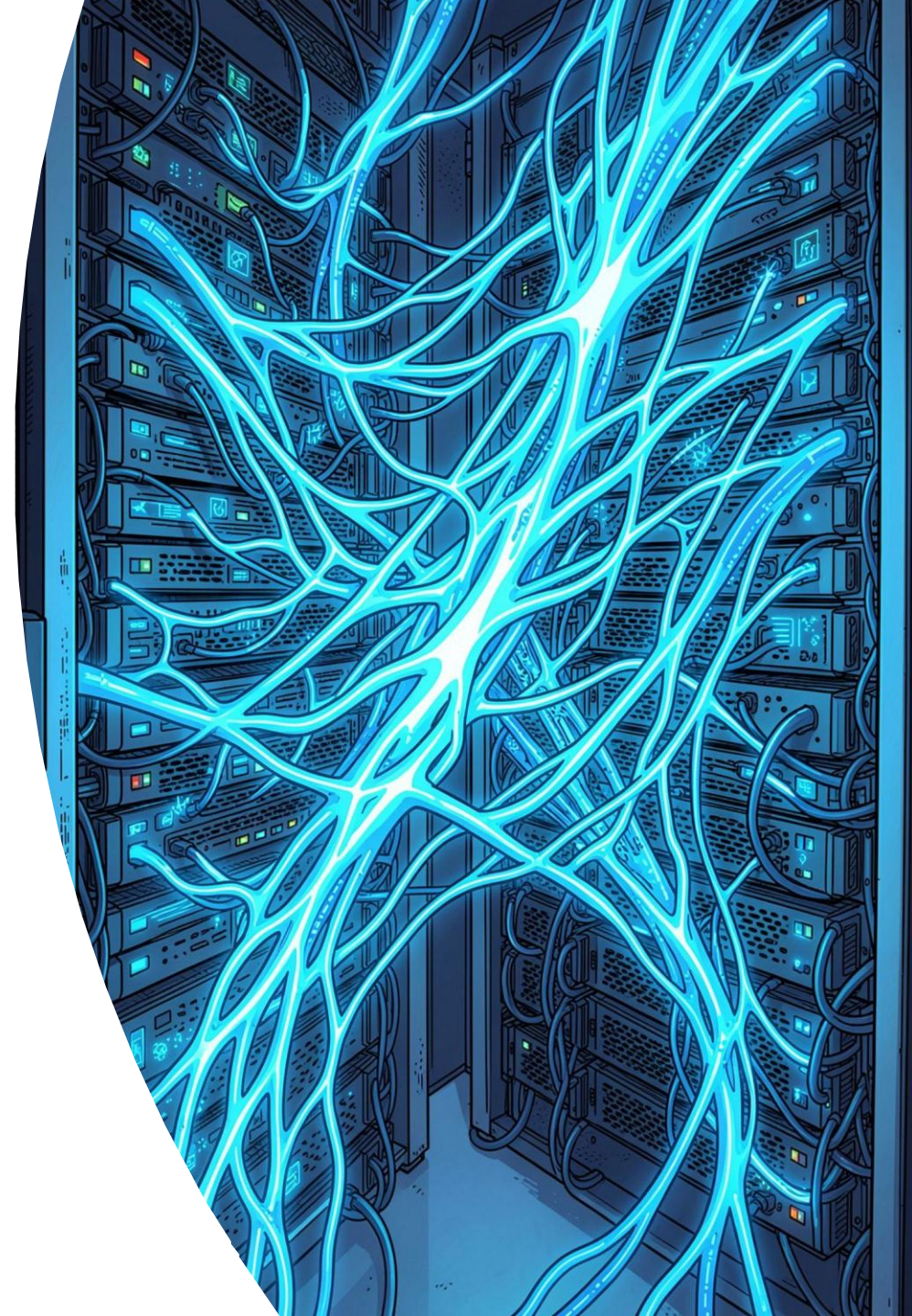
RAG & Enterprise AI Applications

A comprehensive 2-day, 16-hour technical course taking engineers and architects from raw enterprise data to production-grade, secure AI systems — with hands-on labs at every stage.

2 DAYS · 16 HOURS

HANDS-ON LABS

ENTERPRISE FOCUS



Who This Course Is For

Software & AI Engineers

Build and optimize RAG pipelines end-to-end

Data Engineers & Architects

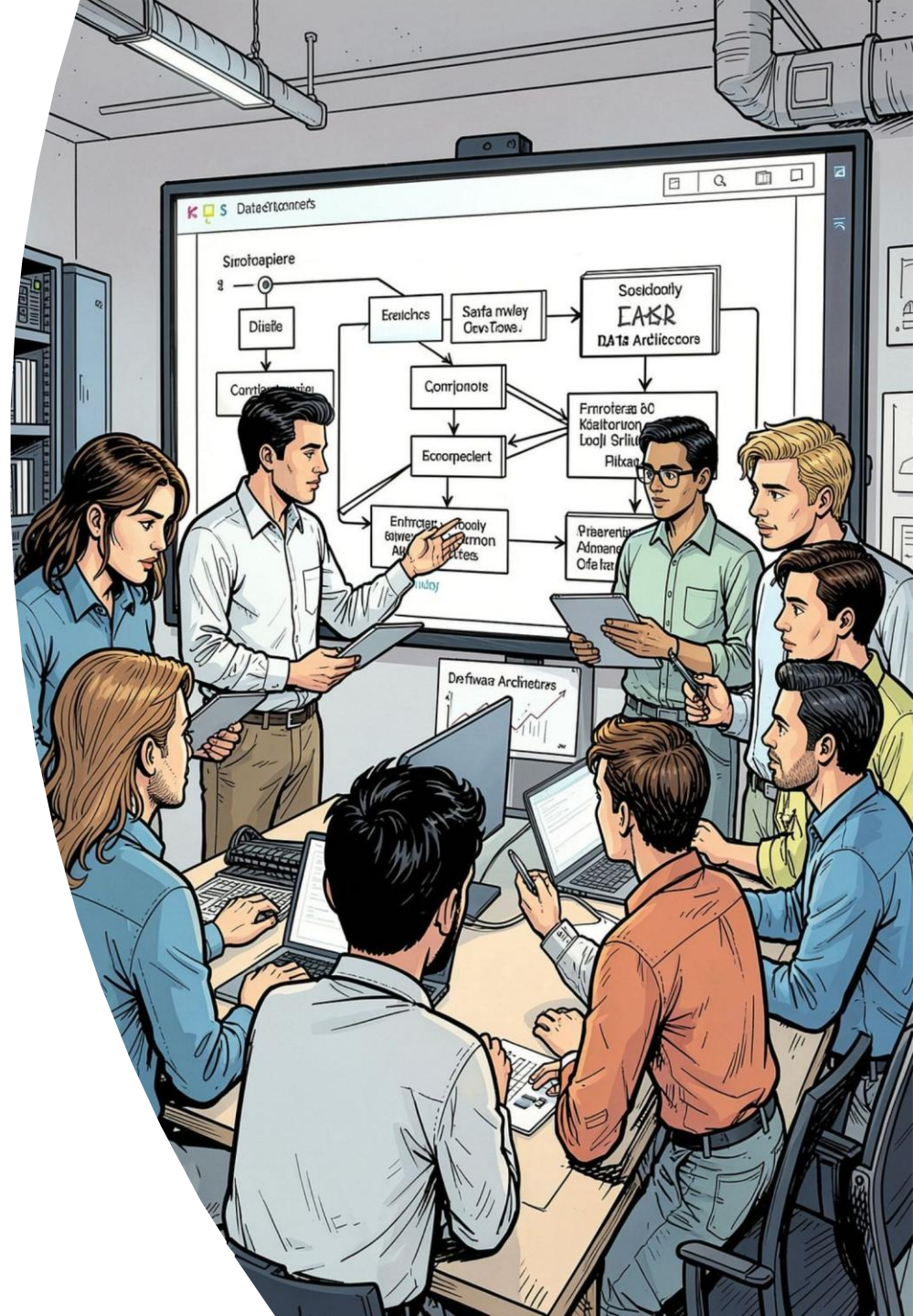
Design ingestion, chunking, and vector storage strategies

Solution Architects

Make architecture decisions for enterprise deployments

Technical Consultants

Guide clients through AI adoption and governance



What You Should Know Before This Course

Recommended baseline skills for the 2-day, 16-hour hands-on RAG program.

Basic Programming

Comfort with Python or a comparable programming language.
Variables, functions, loops, conditionals and basic debugging.

GenAI & LLM Fundamentals

Basic understanding of Generative AI, LLMs and prompts.
Awareness of how LLM applications use external knowledge.

APIs & Application Concepts

Basic understanding of APIs and request/response patterns.
Familiarity with how software applications exchange data.

Data & Database Basics

Comfort working with documents and structured data.
Basic familiarity with databases and data storage concepts.

No prior hands-on RAG experience is required. The course builds from RAG architecture and ingestion through advanced retrieval and production co

The Learning Journey

Two progressive days — from foundational RAG architecture to production-ready enterprise AI.

Day 1 – Foundation

From organizational data to high-quality retrieval

1. Enterprise RAG Architecture
2. Data Ingestion & Preparation
3. Chunking & Embeddings
4. Vector Databases
5. Retrieval Engineering
6. Build a Basic Enterprise RAG

Day 2 – Advanced & Production

From basic RAG to production-ready enterprise AI

1. Advanced RAG Techniques
2. RAG Evaluation & Optimization
3. Security & Governance
4. Enterprise Architecture Patterns
5. Deployment & Observability
6. Capstone Project

Day 1 – Building the Enterprise RAG Foundation

Participants progressively build a working RAG pipeline through guided labs.

01

Enterprise RAG Architecture

RAG lifecycle, patterns, decision frameworks

02

Data Ingestion & Preparation

PDFs, HTML, OCR, metadata extraction

03

Chunking & Embeddings

Semantic, hierarchical chunking; embedding selection

04

Vector Databases

Similarity search, indexing, metadata filtering

05

Retrieval Engineering

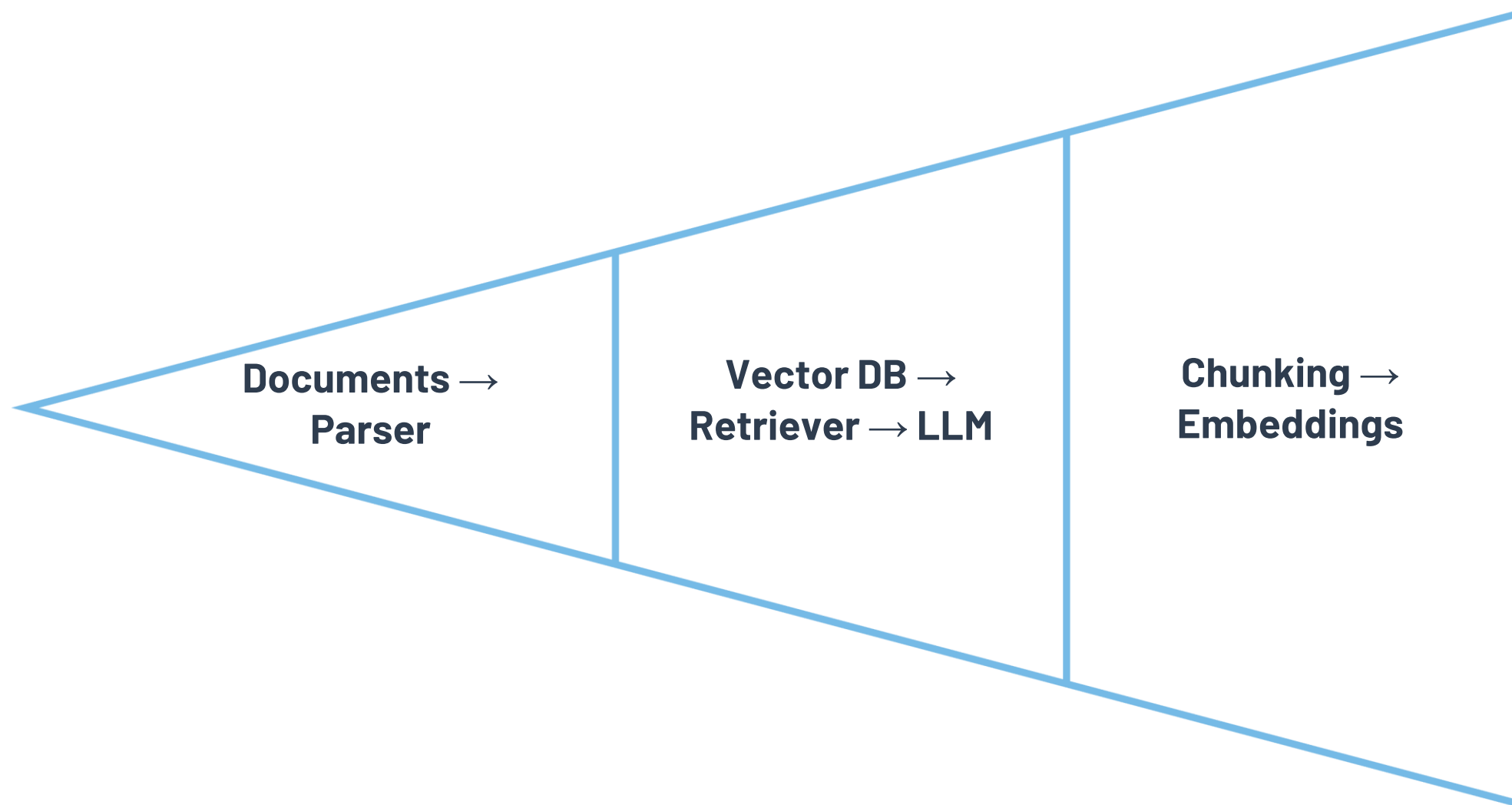
Semantic, BM25, hybrid retrieval; query expansion

06

Baseline RAG Application

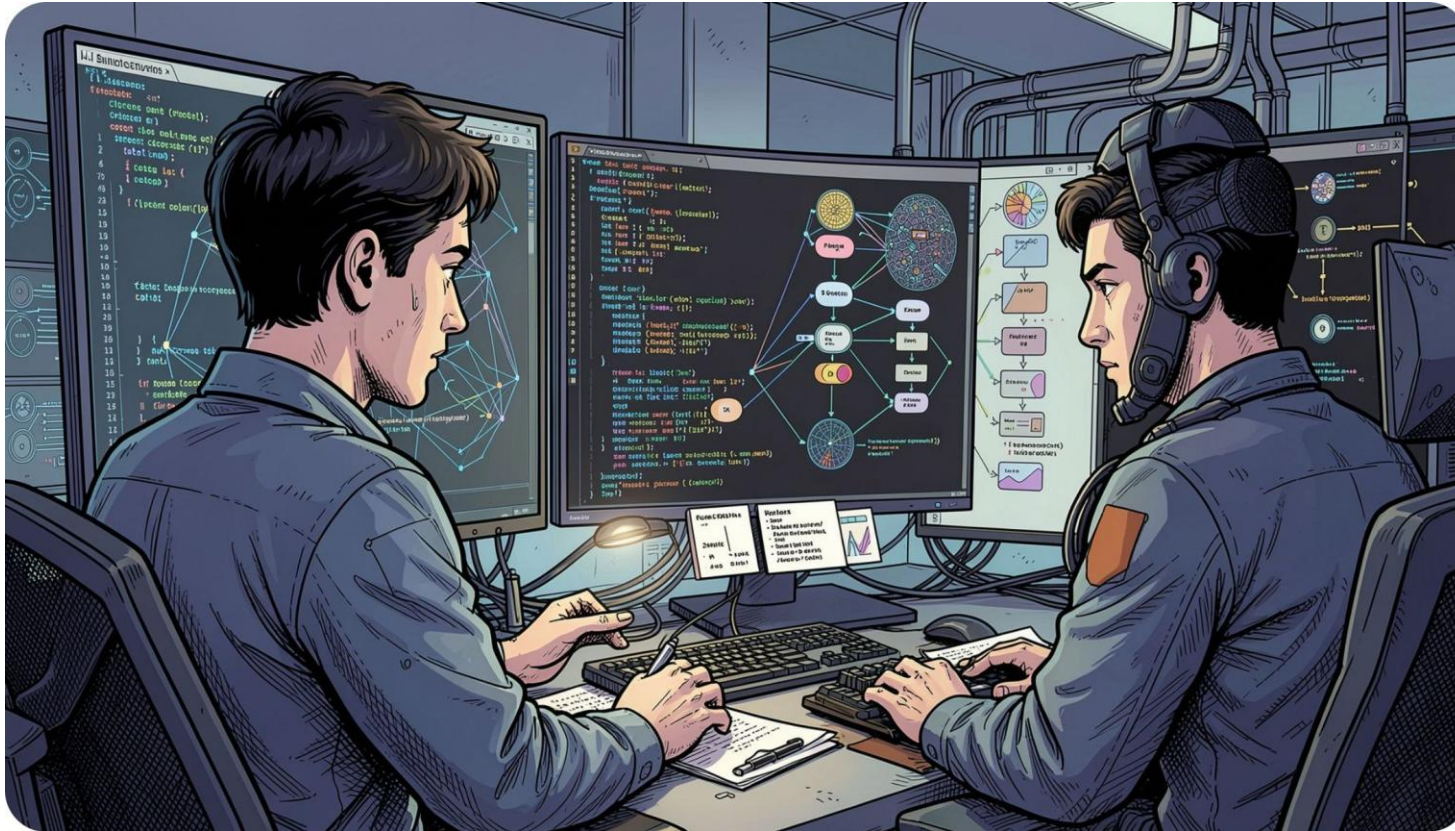
Connect retriever to LLM; citations; grounding

Day 1 – The RAG Pipeline You'll Build



By end of Day 1, participants have a functional baseline RAG application — a working enterprise knowledge assistant with citations and grounded generation.

Day 2 – Advanced RAG & Production Engineering



Advanced Retrieval Techniques

- Reranking & contextual retrieval
- Hierarchical & parent-child retrieval
- Adaptive & agentic RAG
- Query decomposition & multi-query

Evaluation & Optimization

- Retrieval precision/recall, context relevance
- Hallucination detection & faithfulness
- Latency, token usage, cost optimization

Security & Governance

- RBAC/ABAC, document-level permissions
- Prompt injection, data leakage prevention
- Audit trails, data lineage, responsible AI

Enterprise Security & Governance

Treated as a core engineering module — not an optional discussion.

Access Control

Authentication, authorization, RBAC/ABAC, tenant isolation, document-level permissions

AI Security

Prompt injection, indirect injection, data exfiltration, retrieval poisoning, secrets management

Governance & Compliance

Data lineage, audit trails, access logging, model governance, human oversight

Responsible AI

PII handling, sensitive information controls, compliance frameworks, accountability

Enterprise Deployment Architecture

Participants understand the evolution from prototype to production AI platform.

1

Prototype RAG

Local, single-user, experimental

2

Application RAG

API-backed, integrated, tested

3

Enterprise RAG

Multi-tenant, secure, governed

4

Production AI Platform

Scalable, observable, CI/CD

Key areas: REST APIs, containers, cloud deployment, monitoring, logging, tracing, evaluation pipelines, high availability, and cost management.

Capstone Project

What You'll Build

An enterprise-grade RAG application covering the full lifecycle:

- Ingest, chunk, and embed organizational knowledge
- Semantic/hybrid retrieval with citations
- Grounded LLM generation with source attribution
- Evaluation framework with measurable quality metrics
- Security controls and access management
- API interface with monitoring/observability

Suggested Tracks

Digital Engineering

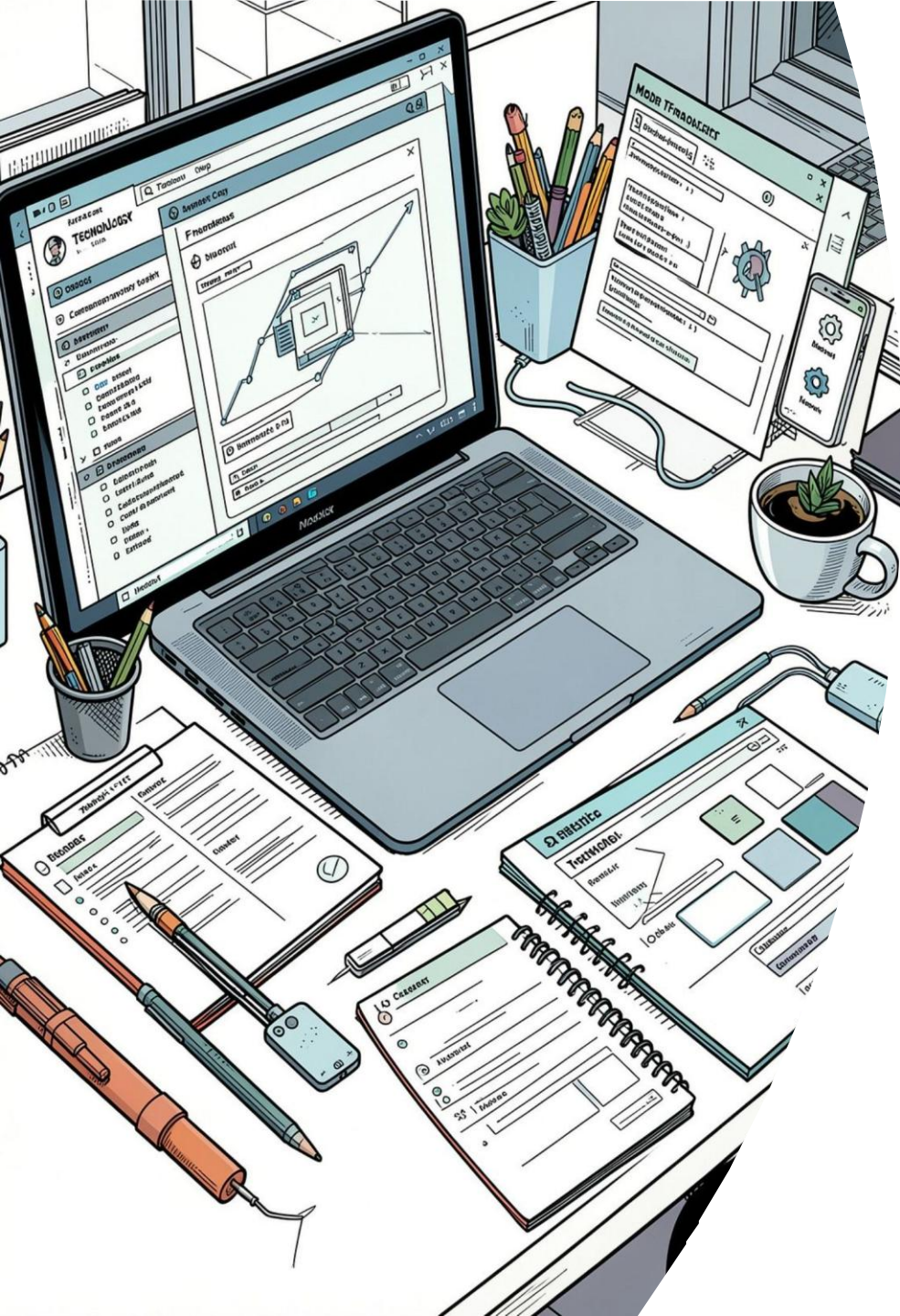
Architecture docs, code repo, IT ops assistant

Data Management

Data catalog, governance, lineage assistant

Business/Industry

Insurance, banking, healthcare, HR, regulatory



Technology Stack

Architecture-first, implementation-practical. Participants work with a real stack while learning transferable patterns.



Core Development

Python · LangChain / LlamaIndex · Open-source or hosted LLM APIs



Vector Databases

FAISS (local) · Chroma (dev) · Qdrant / Weaviate / Milvus (production)



Enterprise Integration

REST APIs · FastAPI · Docker · Cloud services · Observability tools



Embeddings

Open-source embedding models · Commercial embedding APIs where appropriate

By the End of This Course, You Will Be Able To

Participants progress from RAG foundations to an enterprise-oriented implementation approach.

Design Enterprise RAG Architecture

Identify RAG patterns and design an end-to-end architecture for enterprise use cases.

Build RAG Pipelines

Implement ingestion, parsing, chunking, embeddings, vector storage and retrieval.

Engineer Advanced Retrieval

Apply semantic, hybrid, reranking, query transformation and adaptive retrieval approaches.

Evaluate & Optimize RAG

Measure retrieval and answer quality, reduce hallucinations, and optimize latency, tokens and cost.

Apply Security & Governance

Address access control, privacy, security, Responsible AI, auditability and governance.

Deploy & Observe

Understand APIs, containers, monitoring, logging, tracing and production deployment considerations.

Panaashe Certified Enterprise RAG Developer • Awarded upon successful completion of hands-on labs, practical assessment and capstone project.