

PROFESSIONAL MASTERY PROGRAM

GENERATIVE AI & AGENTIC AI ENGINEERING

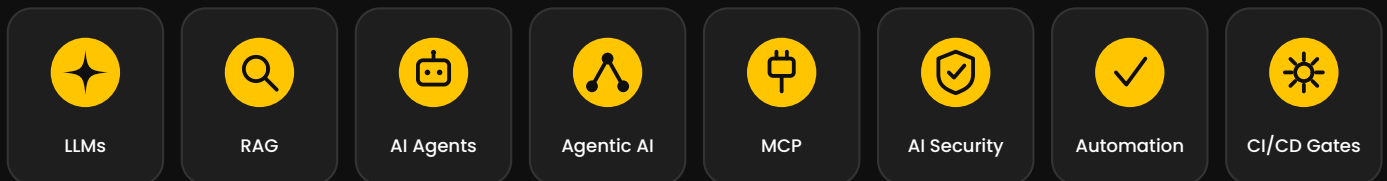
Build · Test · Evaluate · Secure AI

A production-focused learning journey for modern AI Quality Engineers.

LLM Testing · RAG Evaluation · AI Agents · Agentic Workflows · MCP · AI Security · Test Automation



WHAT YOU WILL MASTER



ENROLL TODAY

Start your AI career with RaiseTech

9390271992 | 8247329739

Flat No. 308, Nilgiri Block, Aditya Enclave, Ameerpet, Hyderabad - 500016

OVERVIEW

Build, Test & Secure AI Systems

This program takes learners from software-testing fundamentals to production-ready AI Quality Engineers. It combines classic QA with GenAI evaluation, AI agents, agentic workflows and MCP in a production-focused learning journey.

11

Modules

5

Real-Time Projects

1

Capstone

Career

Preparation

WHAT YOU WILL LEARN

- Explain LLM, RAG, agent, agentic and MCP architectures.
- Design functional, bias, safety, security and performance tests.
- Automate evaluation, UI testing and CI/CD quality gates.

WHO IT IS FOR

- Software / QA Engineers
- SDETs
- Developers
- Data Scientists
- Tech Leads
- AI Product Managers

Focused on reliability, safety and compliance of AI systems.

TECHNOLOGY STACK

TESTING

Python pytest Playwright Git

AI

OpenAI / Azure OpenAI Ollama LangChain LangGraph

DATA & EVALUATION

Chroma Qdrant DeepEval Ragas LangSmith OpenAI Evals

PROTOCOL & SECURITY

MCP MCP Inspector OWASP GenAI Security Project promptfoo

DEVOPS

Docker GitHub Actions

YOUR LEARNING PATH

From Foundations to Capstone



00

Python & Test-Engineering Foundations for AI QA

Python, pytest and QA scripting applied to AI test-data handling.

FOUNDATIONS



01

AI, Generative AI & LLM Foundations

AI/ML/DL, transformers, tokens, temperature and non-determinism.

GENAI & LLM



02

Unique Testing Challenges, Risk & Compliance

Hallucination, bias, privacy risks, EU AI Act and NIST AI RMF.

RISK & COMPLIANCE



03

Building & Functionally Testing LLM Applications

Build an LLM app, then automate its functional test suite.

LLM APPLICATIONS



04

Embeddings, Vector Databases & Retrieval Testing

Validate the retrieval layer: chunking, Top-K and ranking quality.

RETRIEVAL



05

Building & Testing RAG Applications

Build a full RAG pipeline and evaluate it end to end.

RAG



06

Building & Testing AI Agents

Tool-using agents: function calling, planning, memory, tool sequences.

AI AGENTS



07

Agentic AI & Multi-Step Workflow Testing

Multi-step workflows, replanning and human-in-the-loop checkpoints.

AGENTIC AI



08

MCP Foundations, Development & Testing

MCP architecture, server development, schema and security testing.

MCP



09

AI Security, Red Teaming & Responsible AI

OWASP GenAI risks, prompt injection, red teaming and fairness.

AI SECURITY



10

Frontend & End-to-End Testing for AI Applications

Playwright UI automation for chat and agent interfaces.

E2E TESTING



Projects & Capstone

Five real-time projects and the Enterprise AI Quality Engineering Platform.

PROJECTS & CAPSTONE

WHAT YOU WILL LEARN

Curriculum at a Glance

MODULE 00


Python & Test-Engineering Foundations for AI QA

Python Setup · pytest Fixtures · REST API Calls · AI-API Client · Reporting

MODULE 01


AI, Generative AI & LLM Foundations

Transformers · LLMs, SLMs, VLMs · Tokens & Context Window · Temperature & Top-p

MODULE 02


Unique Testing Challenges, Risk & Compliance

Hallucination, Bias & Toxicity · Privacy Risk · EU AI Act · NIST AI RMF

MODULE 03


Building & Functionally Testing LLM Applications

Prompt Templates · JSON Outputs · Faithfulness & PII Leakage · PASS/FAIL Thresholds

MODULE 04


Embeddings, Vector Databases & Retrieval Testing

Embeddings · Similarity Scoring · Chunking Strategy · Top-K & Reranking · Hit Rate / MRR / NDCG

MODULE 05


Building & Testing RAG Applications

Naive vs Advanced RAG · Hybrid Search · Groundedness · Golden Dataset · RAG Regression

MODULE 06


Building & Testing AI Agents

Function Calling · Planning & Memory · Tool-Sequence Testing · 100+ Scenarios

MODULE 07


Agentic AI & Multi-Step Workflow Testing

Multi-Step Workflows · Replanning · Human-Approval Boundary · Excessive Agency

MODULE 08


MCP Foundations, Development & Testing

Host, Client & Server · Tool Schemas · MCP Inspector · Authorization · Tool Poisoning

MODULE 09


AI Security, Red Teaming & Responsible AI

Prompt Injection · Jailbreaks · Red-Team Automation · Attack-Success Rate · Fairness Testing

MODULE 10


Frontend & End-to-End Testing for AI Applications

Chat/Agent UI Automation · Streaming & Async Waits · Visual & Accessibility Checks

5 PROJECTS + CAPSTONE


Projects & Capstone

5 Real-Time Projects · Enterprise AI Quality Engineering Platform · CI/CD Quality Gates · Career Preparation



MODULE 00

Python & Test-Engineering Foundations for AI QA

Python fundamentals, pytest, and QA scripting skills applied specifically to AI test-data handling.

TOPICS

- ▶ Python setup, venv & dependency management
- ▶ Data types, control flow, functions, OOP basics
- ▶ JSON/CSV, REST API calls, logging & exceptions
- ▶ pytest fixtures, parametrization, markers, reporting

HANDS-ON

- ▶ Build reusable Python AI-API client
- ▶ Read prompt/response data from JSON & CSV
- ▶ Create parameterized pytest suites with HTML reports

TOOLS

Python pytest VS Code Git



MODULE 01

AI, Generative AI & LLM Foundations

Conceptual grounding in AI/ML/DL, transformer architecture, and core LLM terminology needed before testing them.

TOPICS

- ▶ AI > ML > DL > Neural Networks > Transformers
- ▶ Foundation models, LLMs, SLMs, VLMs, multimodal & diffusion
- ▶ Tokens, context window, temperature, top-p, prompts
- ▶ Deterministic vs probabilistic systems; why traditional assertions fail

HANDS-ON

- ▶ Call an LLM API in Python
- ▶ Vary temperature/top-p and compare outputs
- ▶ Run the same prompt N times to observe non-determinism

TOOLS

OpenAI / Azure OpenAI API Ollama Python



MODULE 02

Unique Testing Challenges, Risk & Compliance

What makes AI QA different from traditional QA, and the risk/regulatory lens every AI tester must apply.

TOPICS

- ▶ Hallucination, bias, toxicity, privacy risk categories
- ▶ Ethical & legal implications
- ▶ EU AI Act, NIST AI RMF, data-privacy norms
- ▶ Role & career path of the AI Quality Engineer

HANDS-ON

- ▶ Build an AI-application risk & testing checklist
- ▶ Map a sample chatbot against a compliance/risk matrix

TOOLS

Python Documentation templates



MODULE 03

Building & Functionally Testing LLM Applications

Build a small LLM-powered app, then design and automate its functional test suite end to end.

TOPICS

- ▶ LLM architecture & prompt-template design
- ▶ Structured/JSON outputs, error handling, retries, rate limits, timeouts
- ▶ Relevance, correctness & consistency testing
- ▶ Hallucination, faithfulness, toxicity, bias & PII leakage
- ▶ System-prompt leakage, refusal & boundary, multilingual testing

HANDS-ON

- ▶ Build a sample AI assistant
- ▶ Create a 100-prompt test dataset
- ▶ Automate evaluation metrics with PASS/FAIL thresholds

TOOLS

Python pytest DeepEval promptfoo

OpenAI / Azure OpenAI

PROJECT

Project 1 — Automated LLM Quality Evaluation Framework



MODULE 04

Embeddings, Vector Databases & Retrieval Testing

Validate the retrieval layer independently of generation.

TOPICS

- ▶ Embeddings, vector dimensions, cosine/Euclidean similarity
- ▶ Chunking strategy, size/overlap, metadata design
- ▶ Top-K retrieval, similarity scoring, reranking

HANDS-ON

- ▶ Create & store embeddings; run semantic search
- ▶ Inject bad/duplicate chunks and validate detection
- ▶ Test ranking quality: Hit Rate / MRR / NDCG

TOOLS

Chroma / Qdrant

Embedding & reranker models

Python



MODULE 05

Building & Testing RAG Applications

Full RAG pipeline construction and its dedicated evaluation discipline — the most in-demand AI-QA skill today.

TOPICS

- ▶ Indexing, retrieval & generation; naive vs advanced RAG
- ▶ Hybrid search, query transformation, reranking
- ▶ Context precision/recall, groundedness, faithfulness, answer relevance
- ▶ Stale, missing & conflicting document testing; RAG regression

HANDS-ON

- ▶ Build a RAG assistant end to end
- ▶ Create a golden evaluation dataset
- ▶ Inject retrieval defects; compare chunk-size / Top-K configurations

TOOLS

LangChain

Chroma / Qdrant

Ragas

DeepEval

PROJECT

Project 2 — Enterprise RAG Testing & Evaluation Platform



MODULE 06

Building & Testing AI Agents

Move from single-shot LLM calls to tool-using agents, and the new test surface that creates.

TOPICS

- ▶ Agent architecture: tools, function calling, planning, memory
- ▶ Short-term vs long-term memory; observe-plan-act loop
- ▶ Tool selection, argument validation & tool-sequence testing

HANDS-ON

- ▶ Build an IT-support style agent with real tool calls
- ▶ Create 100+ agent test scenarios
- ▶ Inject wrong tool selection / invalid parameters and measure recovery

TOOLS

LangChain

LangGraph

Python

PROJECT

Project 3 — AI Agent Automated Testing Framework



MODULE 07

Agentic AI & Multi-Step Workflow Testing

Goal-driven, multi-step agentic workflows with human-in-the-loop checkpoints — the current frontier of AI QA.

TOPICS

- ▶ AI agent vs. agentic AI; dynamic decision-making & replanning
- ▶ Multi-step workflow & state management; multi-agent basics
- ▶ Human-approval boundary & excessive-agency testing
- ▶ Partial-completion, recovery & agentic regression testing

HANDS-ON

- ▶ Build a multi-step incident-response workflow
- ▶ Inject failed tool calls and test replanning
- ▶ Test human-approval bypass scenarios

TOOLS

LangGraph

Python

pytest

PROJECT

Project 4 — Agentic AI Workflow & Incident-Response Testing Platform



MODULE 08

MCP Foundations, Development & Testing

MCP connects AI systems to tools/data and is a must-have skill for AI QA.

TOPICS

- ▶ MCP architecture: host, client, server, tools, resources, prompts
- ▶ Build/expose an MCP server with tool schemas
- ▶ Tool discovery, schema & input validation
- ▶ Authentication, authorization & unauthorized tool access
- ▶ Malicious tool responses & tool-poisoning concepts

HANDS-ON

- ▶ Build a small MCP server with real tools
- ▶ Debug via MCP Inspector; malformed/missing arguments
- ▶ Automate MCP functional & security suites

TOOLS

MCP

MCP Inspector

pytest

promptfoo

PROJECT

Project 5 — MCP Automated Testing & Security Validation Framework



MODULE 09

AI Security, Red Teaming & Responsible AI

Adversarial testing against OWASP GenAI risks and structured red-team methodology.

TOPICS

- ▶ Prompt injection, jailbreaks
- ▶ System-prompt leakage, PII disclosure, insecure output handling
- ▶ Excessive agency; RAG/knowledge-base/vector-DB/memory poisoning
- ▶ Fairness, bias-score & stereotype-detection testing

HANDS-ON

- ▶ Build attack-prompt dataset & run automated red-team tests
- ▶ Calculate attack-success rate & data-leakage rate
- ▶ Generate a consolidated AI security report

TOOLS

promptfoo

OWASP GenAI Security Project

Python

pytest



MODULE 10

Frontend & End-to-End Testing for AI Applications

AI features ship inside real products — validating the UI layer that wraps the model.

TOPICS

- ▶ UI automation strategy for chat/agent interfaces
- ▶ Streaming responses, async waits & non-deterministic UI states
- ▶ Visual and accessibility checks for AI-driven interfaces

HANDS-ON

- ▶ Automate an AI chatbot UI with Playwright
- ▶ Combine UI checks with backend LLM-quality assertions

TOOLS

Playwright

TypeScript / Python

PUT SKILLS TO WORK

Projects & Outcomes



PROJECT 01

Automated LLM Quality Evaluation Framework

Build a sample AI assistant, create a 100-prompt test dataset and automate evaluation metrics with PASS/FAIL thresholds.

Python

pytest

DeepEval

promptfoo

OpenAI / Azure OpenAI

EXPECTED OUTCOME

Automated LLM quality checks covering functional, hallucination, faithfulness, bias and PII-leakage testing.



PROJECT 02

Enterprise RAG Testing & Evaluation Platform

Build a RAG assistant end to end, create a golden evaluation dataset, inject retrieval defects and compare chunk-size / Top-K configurations.

LangChain

Chroma / Qdrant

Ragas

DeepEval

EXPECTED OUTCOME

Measured context precision/recall, groundedness, faithfulness and answer relevance, with RAG regression testing.



PROJECT 03

AI Agent Automated Testing Framework

Build an IT-support style agent with real tool calls, create 100+ agent test scenarios and inject wrong tool selection or invalid parameters.

LangChain

LangGraph

Python

EXPECTED OUTCOME

Automated tool-selection, argument-validation and tool-sequence tests, with recovery measured.



PROJECT 04

Agentic AI Workflow & Incident-Response Testing Platform

Build a multi-step incident-response workflow, inject failed tool calls to test replanning and test human-approval bypass scenarios.

LangGraph

Python

pytest

EXPECTED OUTCOME

Validated replanning, recovery and human-approval boundaries in agentic workflows.



PROJECT 05

MCP Automated Testing & Security Validation Framework

Build a small MCP server with real tools, debug it via MCP Inspector and automate MCP functional and security suites.

MCP

MCP Inspector

pytest

promptfoo

EXPECTED OUTCOME

Automated MCP functional and security test suites, including malformed-argument handling.



CAPSTONE

Enterprise AI Quality Engineering Platform

Capstone that integrates all five projects into one end-to-end platform across LLM, RAG, agent, agentic, MCP and security layers.

Docker

GitHub Actions

pytest

Python

EXPECTED OUTCOME

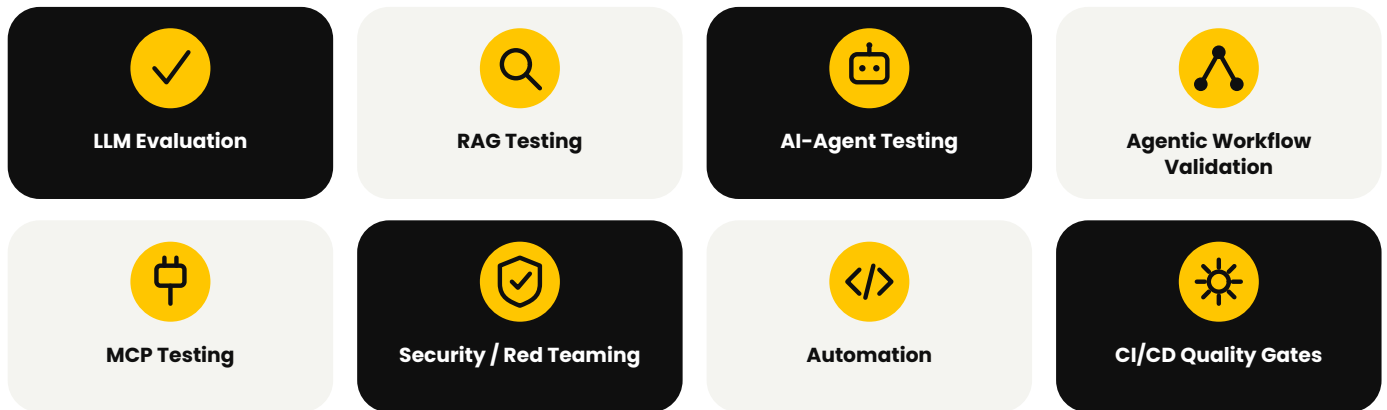
A consolidated platform with CI/CD threshold-based quality gates and a quality dashboard / report.

CAPSTONE OUTCOME

Enterprise AI Quality Engineering Platform

An end-to-end platform that brings together LLM evaluation, RAG testing, AI-agent testing, agentic workflow validation, MCP testing, security/red teaming, automation and CI/CD quality gates.

WHAT THE CAPSTONE INTEGRATES



CAPSTONE DELIVERABLES

- ▶ GitHub repository with architecture diagram and test strategy document
- ▶ AI test cases, golden dataset, and defined evaluation metrics
- ▶ Automated test framework covering LLM, RAG, Agent, Agentic, MCP & Security layers
- ▶ CI/CD pipeline with threshold-based quality gates (Docker + GitHub Actions)
- ▶ Consolidated quality dashboard / report



AI IN YOUR WORKFLOW

The course joins AI systems, evaluation, automation and quality engineering into one workflow: learners build LLM, RAG, agent, agentic and MCP applications, evaluate them with metrics and golden datasets, automate the tests, and secure the result through red teaming and CI/CD quality gates.

BUILD AI SYSTEMS

TEST AI

EVALUATE AI

AUTOMATE AI TESTING

SECURE AI

WHAT YOU WILL GAIN

Skills & Career Preparation

SKILLS YOU WILL GAIN

- Python & pytest Test Automation
- AI & LLM Fundamentals
- AI Risk & Compliance
- LLM Application Testing
- Hallucination & Faithfulness Testing
- Embeddings
- Vector Database Testing
- Retrieval Evaluation
- RAG Evaluation
- AI Agent Testing
- Agentic Workflow Testing
- MCP Development & Testing
- AI Security Testing
- Red Teaming
- Responsible AI & Fairness Testing
- UI / E2E Testing with Playwright
- Test Automation
- CI/CD Quality Gates



ROLES THIS PREPARES YOU FOR

► AI Quality Engineer

Built for Software/QA Engineers, SDETs, Developers, Data Scientists, Tech Leads and AI Product Managers moving into AI quality, safety and compliance.



PRODUCTION-FOCUSED QUALITY ENGINEERING

Explain LLM/RAG/agent/agentic/MCP architectures, design functional, bias, safety, security and performance tests, and automate evaluation, UI testing and CI/CD quality gates.

CAREER PREPARATION



Interview Question Banks

AI / LLM / RAG / Agent / MCP / Security testing interview question banks.



Production Scenarios

Real-world production-scenario walkthroughs and framework-design questions.



Resume, LinkedIn & GitHub

Resume, LinkedIn profile and GitHub portfolio guidance.



Mock Interviews

Mock interview sessions with feedback.



LEARN TODAY, LEAD TOMORROW

We bridge the gap between learning and real-world application with hands-on, project-based training. Join RaiseTech and start building your career in AI Quality Engineering.



CALL US

9390271992



CALL US

8247329739



VISIT US

Flat No. 308, Nilgiri Block, Aditya Enclave, Ameerpet, Hyderabad – 500016

Ready to start your AI career?

Call or visit RaiseTech, Ameerpet to know more about batches and fees.

