ENTERPRISE PROGRAM BRIEF

AI Evaluation and Quality Engineering (O01)

Build a repeatable quality gate for a supplied AI application.

VNode ITeS | O01

DURATION	LEVEL	FORMAT
40 hours	Practitioner / specialist	Private cohort

AUDIENCE PROFILE

Built for these roles

QA engineers, AI engineers, Technical reviewers

OVERVIEW

Executive overview

Build a repeatable quality gate for a supplied AI application. Proposed instructor-led duration; delivery subject to environment and lab validation.

READINESS

Prerequisites

- Can run and modify a working AI application
- Python and API competence
- Applied AI Engineering with Microsoft Foundry, or equivalent demonstrable experience

PROGRAM OUTCOMES

Capabilities your teams will gain

- Implement a versioned evaluation dataset and scorer suite
- Calibrate automated results against references and human review
- Block critical unsafe-action and access regressions
- Defend an evidence-based release or hold recommendation

ENTERPRISE CUSTOMIZATION

Enterprise customization

Tailor this program to your organization's priorities: Build a repeatable quality gate for a supplied AI application.

- Private cohort scheduling
- Equivalent-experience entry review

CURRICULUM

Curriculum roadmap

1

MODULE 1 · 8 HOURS

Evaluation strategy and task contracts

Define quality criteria and failure severity for an enterprise assistant.

- Task success
- Failure taxonomy
- Quality versus safety
- Release thresholds

2

MODULE 2 · 8 HOURS

Datasets and reference evidence

Create normal, ambiguous, denied-access and attack cases with reference evidence.

- Representative cases
- Ground truth
- Adversarial cases
- Dataset versioning

3

MODULE 3 · 8 HOURS

Scorers and human calibration

Implement checks and calibrate automated scoring against human labels.

- Deterministic tests
- LLM judges
- Human review
- Bias and judge disagreement

4

MODULE 4 · 8 HOURS

Agent and retrieval diagnostics

Find root causes and compare baseline and changed application versions.

- Tool correctness
- Retrieval quality
- Trace analysis
- Cost and latency

5

MODULE 5 · 8 HOURS

Quality engineering capstone

Run the evaluation gate and defend a release or hold recommendation.

- CI integration
- Threshold enforcement
- Exceptions
- Release review

PREPARATION

Environment prerequisites

Isolated Azure subscription/resource group Scoped deployment rights and resource-provider access Approved region, quota and optional model deployment Cost budget and cleanup plan Delivery depends on environment and lab validation. This outline does not include step-by-step lab instructions.

CAPSTONE

Capstone

Run the evaluation gate and defend a release or hold recommendation. Assessed AI quality gate Evidence of relevant access, quality and risk controls Demonstration and limitations record

ASSESSMENT

Assessment

Scenario-based practical assessment Implement a versioned evaluation dataset and scorer suite 30% Calibrate automated results against references and human review 25% Block critical unsafe-action and access regressions 25% Defend an evidence-based release or hold recommendation 20% Proposed VNode outline-level assessment; no vendor accreditation or completed lab validation implied.

DELIVERABLES

Participant deliverables

- Assessed AI quality gate
- Evidence of relevant access, quality and risk controls
- Demonstration and limitations record

SCOPE

Scope boundaries

Step-by-step lab instructions; Runnable solution code; Packaged source datasets; Executed environment validation

DELIVERY

Delivery options

Count teaching hours, not calendar days. Exclude breaks. Split into 4-hour or 8-hour teaching sessions; schedule complex project work in contiguous blocks. Proposed hours exclude breaks and are subject to pilot validation.

PRIVATE DELIVERY

Plan a private cohort for your team

Scope delivery format, role mix, and rollout timeline with our solutions team.

Plan Private Delivery

Book advisory call