

AI Testing Industrialization

COURSE CONTENT

GET IN TOUCH



Multisoft Systems
B - 125, Sector - 2, Noida



(+91) 9810-306-956



info@multosoftsystems.com



www.multosoftsystems.com

About Multisoft

Train yourself with the best and develop valuable in-demand skills with Multisoft Systems. A leading certification training provider, Multisoft collaborates with leading technologies to bring world-class one-on-one and certification training. Intending to empower professionals and businesses across the globe, we offer more than 1500 training courses, which are delivered by Multisoft's global subject matter experts. We offer tailored corporate training, a comprehensive learning solution with lifetime e-learning access, after-training support and globally recognized training certificates.

About Course

AI Testing Industrialization is an advanced quality engineering course focused on making AI validation systematic, automated, repeatable, and scalable. It addresses the distinctive testing requirements of machine learning models, generative AI applications, LLMs, RAG solutions, datasets, and production AI services.

The course connects AI testing with automation, CI/CD, MLOps, observability, responsible AI, and governance practices. Its emphasis on continuous evaluation and enterprise testing standards enables participants to understand how AI quality can be managed throughout development, deployment, and production operations rather than treated as a one-time pre-release activity.

Module 1: AI Testing & Industrialization Foundations

- ✓ Introduction to AI quality engineering
- ✓ Evolution from conventional software testing to AI testing
- ✓ Characteristics of deterministic and probabilistic systems
- ✓ AI/ML lifecycle and associated quality risks
- ✓ Testing challenges across AI-enabled applications
- ✓ Concept of AI testing industrialization
- ✓ From experimental validation to repeatable testing practices
- ✓ Quality dimensions for production AI
- ✓ Enterprise AI testing lifecycle
- ✓ Roles and responsibilities within AI quality engineering teams

Module 2: AI/ML System Architecture for Testing

- ✓ Understanding AI/ML solution architecture
- ✓ Data ingestion and preprocessing layers
- ✓ Feature engineering and feature pipelines
- ✓ Training and inference architecture
- ✓ Model serving and prediction APIs
- ✓ Batch versus real-time inference
- ✓ AI application integration patterns
- ✓ Testing dependencies across the AI technology stack
- ✓ Identifying test points and observability requirements
- ✓ Designing testable AI architectures

Module 3: AI Test Strategy & Quality Engineering

- ✓ Defining an AI testing strategy
- ✓ Translating business requirements into AI quality criteria
- ✓ Functional and non-functional AI testing
- ✓ Establishing measurable acceptance criteria
- ✓ Model-level versus application-level validation
- ✓ Risk-based testing for AI solutions
- ✓ Test scenario and test case design
- ✓ Baseline and benchmark definition
- ✓ Evaluation metrics and thresholds
- ✓ Regression strategy for evolving AI systems
- ✓ Quality gates across development and deployment stages
- ✓ AI testing traceability and reporting

Module 4: Data Quality, Dataset Validation & Test Data Management

- ✓ Importance of data quality in AI testing
- ✓ Training, validation and test datasets
- ✓ Data profiling and statistical validation
- ✓ Schema and constraint validation
- ✓ Missing, inconsistent and duplicate data detection
- ✓ Outlier and anomaly analysis
- ✓ Data distribution validation
- ✓ Dataset representativeness
- ✓ Data leakage detection

- ✓ Class imbalance considerations
- ✓ Dataset versioning and lineage
- ✓ Synthetic test data considerations
- ✓ Test dataset management
- ✓ Data quality automation
- ✓ Data quality gates within AI pipelines

Module 5: ML Model Testing & Performance Validation

- ✓ Model testing lifecycle
- ✓ Validation of classification models
- ✓ Validation of regression models
- ✓ Precision, recall, F1-score and accuracy
- ✓ ROC-AUC and confusion matrix analysis
- ✓ Error analysis and failure categorization
- ✓ Threshold testing and calibration
- ✓ Model robustness testing
- ✓ Boundary and edge-case validation
- ✓ Adversarial input considerations
- ✓ Model reproducibility testing
- ✓ Comparing candidate and baseline models
- ✓ Model regression testing
- ✓ Performance and latency validation
- ✓ Establishing model release criteria

Module 6: Generative AI & LLM Testing

- ✓ Quality challenges in generative AI applications
- ✓ Understanding non-deterministic LLM responses
- ✓ Prompt and response testing
- ✓ Creating representative evaluation datasets
- ✓ Golden datasets and reference responses
- ✓ Answer relevance and correctness evaluation
- ✓ Groundedness and faithfulness testing
- ✓ Hallucination detection and evaluation
- ✓ Context relevance testing
- ✓ Semantic similarity evaluation
- ✓ RAG pipeline testing
- ✓ Retrieval quality and ranking evaluation
- ✓ Testing prompt templates and prompt versions
- ✓ Multi-turn conversational testing
- ✓ LLM-as-a-judge approaches
- ✓ Human evaluation workflows
- ✓ Toxicity and harmful-content evaluation
- ✓ Bias and fairness considerations
- ✓ Jailbreak and prompt-injection testing
- ✓ Latency, token usage and cost evaluation
- ✓ Regression testing for LLM applications

Module 7: AI Test Automation Frameworks & Tooling

- ✓ Architecture of an AI testing framework

- ✓ Designing reusable testing components
- ✓ Automated dataset validation
- ✓ Automated model evaluation
- ✓ Automating LLM evaluation workflows
- ✓ Metric calculation and threshold enforcement
- ✓ Test configuration management
- ✓ Parameterized AI test scenarios
- ✓ API-based AI testing
- ✓ Batch evaluation workflows
- ✓ Managing evaluation datasets
- ✓ Automated comparison of model versions
- ✓ Test result storage and analysis
- ✓ Evaluation dashboards and reporting
- ✓ Integration with defect and issue management
- ✓ Framework extensibility for new AI use cases

Module 8: CI/CD, MLOps & Continuous AI Testing

- ✓ Role of testing within MLOps
- ✓ Continuous integration for AI applications
- ✓ Continuous training and continuous evaluation concepts
- ✓ Integrating automated tests into CI/CD pipelines
- ✓ Data validation during pipeline execution
- ✓ Model validation before deployment
- ✓ LLM evaluation within release pipelines
- ✓ Automated quality gates
- ✓ Model registry and version control integration

- ✓ Artifact and experiment traceability
- ✓ Environment-specific validation
- ✓ Canary and controlled deployment validation
- ✓ Rollback criteria
- ✓ Continuous regression testing
- ✓ Release readiness and approval workflows

Module 9: Production Monitoring, Drift & Reliability Testing

- ✓ Transitioning from pre-production testing to production assurance
- ✓ AI observability principles
- ✓ Monitoring prediction quality
- ✓ Data drift detection
- ✓ Concept drift considerations
- ✓ Model performance degradation
- ✓ LLM response quality monitoring
- ✓ Monitoring hallucination and failure patterns
- ✓ Latency and throughput monitoring
- ✓ Availability and reliability indicators
- ✓ Cost and resource utilization monitoring
- ✓ Alert thresholds and escalation
- ✓ Production feedback loops
- ✓ Triggering retraining or reevaluation
- ✓ Incident analysis for AI systems
- ✓ Continuous production quality assessment

Module 10: Responsible AI, Security, Compliance & Governance Testing

- ✓ Responsible AI testing principles
- ✓ Fairness and bias assessment
- ✓ Explainability validation
- ✓ Transparency and traceability
- ✓ Privacy considerations in AI testing
- ✓ Sensitive information exposure testing
- ✓ Security risks in AI applications
- ✓ Prompt injection and malicious-input testing
- ✓ Model and API security considerations
- ✓ AI risk classification
- ✓ Evaluation evidence and audit trails
- ✓ Model documentation and test documentation
- ✓ Approval and governance checkpoints
- ✓ Regulatory and organizational compliance considerations
- ✓ Integrating responsible AI controls into automated testing

Module 11: Enterprise AI Testing Industrialization & Capstone Implementation

- ✓ Designing an enterprise AI quality operating model
- ✓ Standardizing AI testing processes
- ✓ Creating reusable test assets and evaluation libraries
- ✓ Establishing enterprise evaluation datasets
- ✓ Defining organization-wide AI quality metrics

- ✓ Testing standards and quality policies
- ✓ Centralized versus federated testing models
- ✓ AI testing maturity assessment
- ✓ Scaling testing across multiple models and applications
- ✓ Test environment and infrastructure strategy
- ✓ Quality dashboards and executive reporting
- ✓ Defect and failure taxonomy
- ✓ Governance and ownership models
- ✓ Continuous improvement of AI testing practices
- ✓ Capstone: designing an end-to-end AI testing industrialization framework
- ✓ Defining data, model and application-level tests
- ✓ Integrating automated evaluations with CI/CD
- ✓ Establishing production monitoring and quality gates
- ✓ Presenting the final AI quality engineering strategy