Datastat Research
Building Capacity in Training and Research

DATASTAT TRAINING IN

OFFICIAL COURSE PROSPECTUS

Center for Advanced Research & Professional Development

Nairobi Training Center • Global Virtual Academy • Corporate On-Site

ARTIFICIAL INTELLIGENCE AND BLOCK CHAIN

AI Agent Evaluation and Testing Training Course

Comprehensive Course Syllabus, Practical Modules & Upcoming Schedules

PROGRAM SPECIFICATIONS & HIGHLIGHTS

DURATION
5 Days Practical Hands-on

DELIVERY MODES
In-Person Class & Live Virtual

CERTIFICATION
Accredited Certificate of Completion

INTAKES / DATES
Multiple Cohort Schedules Included

PREPARED FOR:

Valued Participant

Email: training@datastatresearch.org | Issued: November 21, 2026

Academic & Professional Training Year: 2026

DATASTAT TRAINING INSTITUTE

P.O. Box 102802-00101, Nairobi, Kenya

Training Centers: Nairobi • Mombasa • Kigali • Dubai • Live Virtual Academy

Tel: +254 724 527 104 | +254 734 969 612

Email: training@datastatresearch.org

www.datastatresearch.org

AI Agent Evaluation and Testing Training Course

Artificial Intelligence And Block Chain • 5 Days Intensive • Accredited Certification

Course Overview

AI Agent Evaluation and Testing Training Course provides a comprehensive framework for designing, validating, benchmarking, and optimizing intelligent AI agent systems in modern enterprise environments.

Key Skills & Software Tools Covered

Agent evaluation • AI testing • Performance benchmarking • Safety validation

Upcoming Training Schedules

Confirmed cohort intakes and venues. Both physical in-person classroom and live virtual online seats are available.

TRAINING DATES	LOCATION / VENUE	FORMAT / DURATION	INVESTMENT (FEE)
21 Sep 2026 – 25 Sep 2026	Nairobi	Classroom / 5d	**$1,500 USD / KES 110,000**
21 Sep 2026 – 25 Sep 2026	Kigali	Classroom / 5d	**$3,000 USD / KES 110,000**
21 Sep 2026 – 25 Sep 2026	Dubai	Classroom / 5d	**$5,000 USD / KES 110,000**
21 Sep 2026 – 25 Sep 2026	Cape Town	Classroom / 5d	**$4,500 USD / KES 110,000**
21 Sep 2026 – 25 Sep 2026	Addis Ababa	Classroom / 5d	**$3,000 USD / KES 110,000**
21 Sep 2026 – 25 Sep 2026	Singapore	Classroom / 5d	**$5,000 USD / KES 110,000**
21 Sep 2026 – 25 Sep 2026	Istanbul	Classroom / 5d	**$5,000 USD / KES 110,000**
21 Sep 2026 – 25 Sep 2026	Bangkok	Classroom / 5d	**$5,000 USD / KES 110,000**
21 Sep 2026 – 25 Sep 2026	Accra	Classroom / 5d	**$4,000 USD / KES 110,000**
28 Sep 2026 – 2 Oct 2026	Nairobi	Classroom / 5d	**$1,500 USD / KES 110,000**
28 Sep 2026 – 2 Oct 2026	Kigali	Classroom / 5d	**$3,000 USD / KES 110,000**
28 Sep 2026 – 2 Oct 2026	Dubai	Classroom / 5d	**$5,000 USD / KES 110,000**

* Corporate discounts and customized cohort scheduling are available for teams of 3 or more participants.

Course Curriculum & Modules

AI Agent Evaluation and Testing Training Course

Introduction

AI Agent Evaluation and Testing Training Course provides a comprehensive framework for designing, validating, benchmarking, and optimizing intelligent AI agent systems in modern enterprise environments. As organizations rapidly adopt Generative AI, Large Language Models (LLMs), Autonomous AI Agents, Agentic Workflows, Retrieval-Augmented Generation (RAG), and AI Automation, the ability to measure agent performance, reliability, safety, and effectiveness has become a critical capability. This course equips professionals with advanced techniques for AI agent testing, evaluation frameworks, quality assurance, prompt evaluation, model benchmarking, hallucination detection, response accuracy measurement, and continuous AI improvement.

Participants will learn how to build robust AI evaluation pipelines using industry best practices covering functional testing, behavioral testing, security validation, performance assessment, observability metrics, human-in-the-loop evaluation, synthetic test generation, and automated benchmarking. Through practical labs and real-world case studies, learners will explore how enterprises evaluate AI agents deployed in customer service, cybersecurity, finance, healthcare, software engineering, and business operations while ensuring trustworthy AI, responsible AI governance, scalability, and production readiness.

Course Duration

5 Days

Course Objectives

By the end of this course, participants will be able to:

- Understand advanced AI Agent Evaluation Frameworks and testing methodologies.
- Design comprehensive AI Agent Testing Strategies for enterprise applications.
- Evaluate LLM-powered agents using accuracy, reliability, and quality metrics.
- Implement automated AI testing pipelines for continuous validation.
- Measure AI agent performance using benchmarking and evaluation metrics.
- Identify and reduce AI hallucinations, bias, and incorrect responses.
- Apply Prompt Engineering Evaluation Techniques for optimized AI behavior.
- Perform Functional, Regression, and End-to-End AI Agent Testing.
- Implement Human-in-the-Loop Evaluation processes.
- Use AI Observability and Monitoring Tools for agent assessment.
- Evaluate Agentic Workflow Performance and Decision-Making Quality.
- Apply Responsible AI Testing and Governance Practices.
- Build enterprise-ready AI Quality Assurance and Validation Frameworks.

Target Audience

- AI Engineers and Machine Learning Engineers
- Data Scientists and AI Researchers
- Software Test Engineers and QA Professionals
- AI Product Managers and Product Owners
- DevOps, MLOps, and AI Platform Engineers
- Enterprise Architects and Solution Architects
- Cybersecurity Professionals working with AI Systems
- Business Analysts and Digital Transformation Leaders

Course Modules

Module 1: Fundamentals of AI Agent Evaluation and Testing

- Introduction to AI Agent Architectures and Evaluation Challenges
- Understanding AI agent lifecycle testing
- Differences between traditional software testing and AI testing
- AI quality dimensions: accuracy, reliability, safety, and usability
- Establishing AI evaluation objectives and success criteria
- Case Study: Evaluating an enterprise customer support AI agent before production deployment.

Module 2: AI Agent Testing Frameworks and Methodologies

- Designing AI testing strategies and test plans
- Functional testing for AI agent capabilities
- Behavioral and conversational testing approaches
- Regression testing for evolving AI models
- Automated testing frameworks for AI applications
- Case Study: Testing an AI banking assistant handling thousands of customer interactions.

Module 3: LLM Evaluation and Benchmarking Techniques

- Understanding LLM evaluation metrics
- Measuring response accuracy and relevance
- Benchmarking different AI models and agents
- Evaluating reasoning and decision-making capabilities
- Using synthetic datasets for AI testing
- Case Study: Comparing multiple LLM-based agents for enterprise knowledge management.

Module 4: Prompt Evaluation and AI Response Quality Testing

- Advanced prompt testing methodologies
- Measuring prompt effectiveness and consistency
- Detecting hallucinations and factual inaccuracies
- Evaluating AI-generated content quality
- Optimizing prompts through iterative testing
- Case Study: Improving an AI research assistant by testing thousands of prompt variations.

Module 5: AI Agent Performance, Reliability, and Scalability Testing

- Load testing AI agents in production environments
- Measuring latency and response performance
- Evaluating multi-agent system scalability
- Stress testing AI workflows
- Monitoring reliability and availability metrics
- Case Study: Performance testing an AI virtual assistant serving millions of users.

Module 6: AI Safety, Security, and Responsible AI Evaluation

- Testing AI agents against harmful behaviors
- Evaluating bias and fairness risks
- Security testing for AI vulnerabilities
- Testing prompt injection and adversarial attacks
- Implementing Responsible AI evaluation frameworks
- Case Study: Security evaluation of an AI cybersecurity monitoring agent.

Module 7: AI Observability, Monitoring, and Continuous Evaluation

- Building AI agent observability frameworks
- Monitoring AI behavior after deployment
- Tracking agent performance metrics
- Continuous evaluation and improvement pipelines
- Integrating AI monitoring platforms
- Case Study: Continuous monitoring of an AI financial advisory agent.

Module 8: Enterprise AI Testing Automation and Future Trends

- Building automated AI evaluation pipelines
- Integrating AI testing with DevOps and MLOps
- Creating AI quality governance models
- Managing AI agent updates and version testing
- Exploring future trends in autonomous AI evaluation
- Case Study: Implementing an enterprise AI testing center of excellence.

Training Methodology

- Interactive lectures and presentations.
- Group discussions and brainstorming sessions.
- Hands-on exercises using real-world datasets.
- Role-playing and scenario-based simulations.
- Analysis of case studies to bridge theory and practice.
- Peer-to-peer learning and networking.
- Expert-led Q&A sessions.
- Continuous feedback and personalized guidance.

Register as a group from 3 participants for a Discount

Send us an email: info@datastatresearch.org or call +254724527104

Certification

Upon successful completion of this training, participants will be issued with a globally- recognized certificate.

Tailor-Made Course

We also offer tailor-made courses based on your needs.

Key Notes

a. The participant must be conversant with English.

b. Upon completion of training the participant will be issued with an Authorized Training Certificate

c. Course duration is flexible and the contents can be modified to fit any number of days.

d. The course fee includes facilitation training materials, 2 coffee breaks, buffet lunch and A Certificate upon successful completion of Training.

e. One-year post-training support Consultation and Coaching provided after the course.

f. Payment should be done at least a week before commence of the training, to DATASTAT CONSULTANCY LTD account, as indicated in the invoice so as to enable us prepare better for you.

Certification & Learning Methodology

- Hands-on practical delivery with real-world datasets, case studies, and lab exercises.
- Post-training mentorship, access to class resources, and customized follow-up support.
- Participants who successfully complete the course receive an accredited Datastat Certificate of Completion.

How to Enroll & Reserve Your Seat

1. Online Registration: Visit https://datastatresearch.org/courses/ai-agent-evaluation-and-testing-training-course to register online.
2. Email Booking: Send participant details to training@datastatresearch.org for official invoice & reservation.
3. Support & Assistance: Inquire via phone +254 724 527 104 | +254 734 969 612.