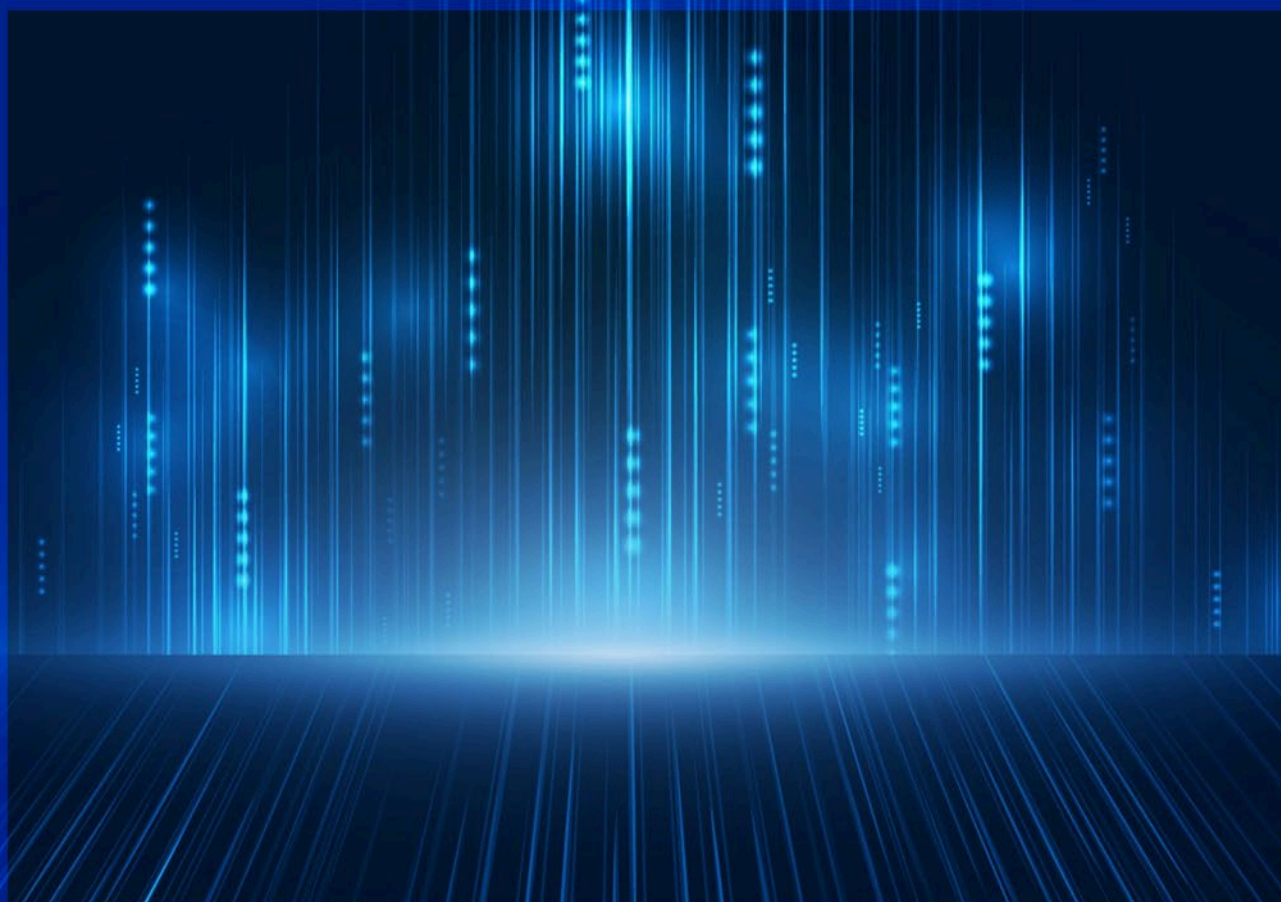


AI Testing and Assurance Services

A KPMG Service for G-Cloud 15



Service Definition



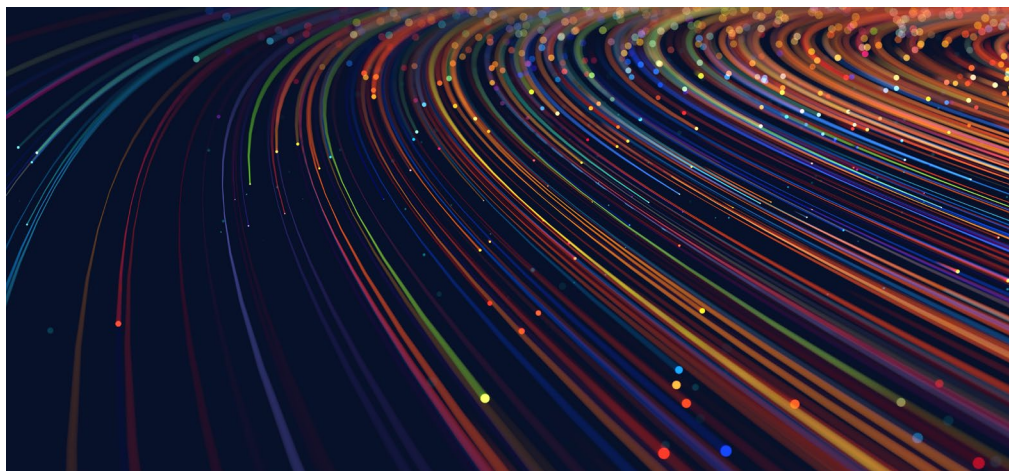
Service Description:

This service is valuable for clients delivering AI solutions into the enterprise.

Testing of AI-powered applications introduces unique challenges due to their dynamic nature, dependency on data, evolving behaviour, and integration with traditional software components.

Ensuring these systems function accurately, ethically, and securely while adhering to evolving AI regulatory standards requires a dynamic adaptive yet robust test approach.

Our expert teams apply proven assurance and testing techniques and apply the latest approach to verification of AI solutions, to provide clients with demonstrable validation that AI solutions are fit for purpose and will meet risk acceptance criteria.



What are the benefits of the service?

- Reduce risk and cost, accelerate AI solution delivery
- Improve quality, accelerate safe AI delivery
- Clear, unbiased insights on AI functionality and performance
- Evaluate AI solutions using manual and automated testing
- Assess adherence to good practice (e.g. EU AI Act)
- Reassure stakeholders on AI risk and good practice
- Standardised, repeatable, and efficient test processes
- Reduce delivery risk by identifying early defects
- Improve final product quality via increased test coverage



Our service features

- Apply KPMG's proven methodology
- Structured, risk-based quality assurance for AI solutions
- Deploy mixed teams: SMEs, AI specialists, test professionals
- Use AI Test Accelerators for productivity and quality
- "Shift-left" testing: early defect identification and resolution
- Increase test coverage with data-driven automation
- Functional and non-functional testing
- AI-specific testing: Model Evaluation, Reliability Testing, Token Usage, Explainability Ratings, Drift Threshold Testing

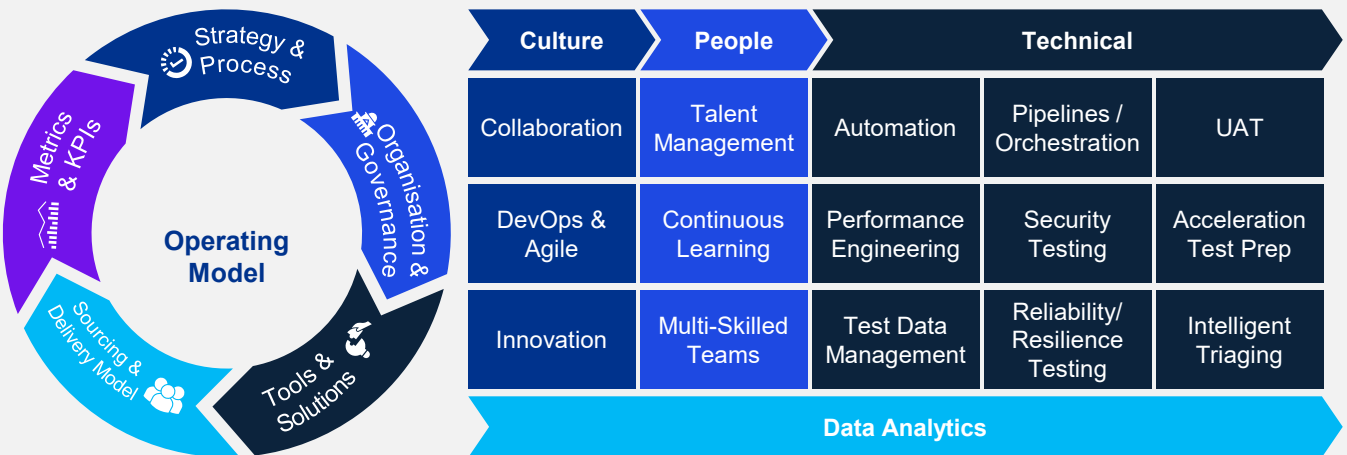
KPMG AI Testing and Assurance Framework

Our model includes assessments that evaluate how effectively our clients test their own AI systems; uncovering gaps in strategy, data, and governance to ensure reliability and compliance.

We also provide a service to take on the end-to-end testing of AI solutions, combining specialist skills, proven frameworks, and continuous monitoring to keep complex models trustworthy in production.

AI Testing Assessments

KPMG's AI Testing Assessment helps customers understand how effective their testing of AI is and what types of risks to address in their AI transformation.



Considerations for Assessment and Readiness for Testing AI Systems

- What types of AI are being deployed?
- Symbolic (rules based; decision trees, use for expert systems)
- Connectionism (neural nets, back-propagation; "learning" from data)
- Evolutionary (candidate models, evolving structures)
- Bayesian (probabilistic classification)
- Analogizing (comparison and association)
- Generative vs Agentic AI solutions

Risk based approach to Testing AI Systems and Solutions

- Algorithmic obscurity
- Radical Fragility
- Validity/veracity of claims/requirements
- Regulatory/liability implications
- Oracle elasticity
- Total cost of Quality
- LLM Mentalist Syndrome (anthropomorphism)

KPMG AI Testing and Assurance Framework

Testing of AI

Aligned to KPMG's Trusted AI Framework, we integrate traditional software testing practices with specialised testing methods for AI-powered applications and components.

	Test Layer	Objective of Test
AI Model Test	Human Evaluation	Evaluate prompt and response coverage against human judgement with context sensitivity and domain expertise
	Automated Evaluation	Evaluate response scores with industry accepted standards and tools (e.g. DeepEval, RAS, ROUGE, F1)
	LLM-as-a-judge	Grading the responses using rubric-based evaluation for tone, accuracy, etc. and rating the response across prompt sets
AI Application Test	Unit Test	Business flow validation of all key features and components and establishing metrics to evaluation dashboard
	API / Integration Test	Validation of responses in and out of integration points work as expected
	Functional / UAT Test	Verify the workflows triggered by AI responses and intended output
	Privacy & Security Test	Validation of PII and regulatory standards data leakage by LLM or vulnerability to prompt injection
	Usability & Accessibility Test	End user perception and usage of the application for the intended purpose adhering to accessibility WCAG 2.1 guidelines
	Performance & Resilience Test	Validation of context window, API rate limit, concurrent user and concurrent prompt tests

LLM as a judge, Human-in-the-loop, and Model Evaluation, utilising a range of AI evaluation metrics applied according to AI application type, ensuring objective and relevant performance assessment.



kpmg.com/uk

© 2026 KPMG LLP, a UK limited liability partnership and a member firm of the KPMG global organisation of independent member firms affiliated with KPMG International Limited, a private English company limited by guarantee. All rights reserved. Document Classification: KPMG Confidential

Document Classification: KPMG Confidential

CREATE. | CRT164925A January 2026