



ITNextGen Limited
Digital Technology Specialists

G-Cloud 15 Service Design Document

Generative AI Testing & Assurance Service

ITNextGen Limited

E: publicsector@itnextgen.co.uk | M: 07545809681 | W: www.itnextgen.co.uk



Service Overview

ITNextGen's Generative AI Testing & Assurance consultants help buyers validate and assure Generative AI and LLM solutions such as chatbots, knowledge assistants (RAG), summarisation, content generation and agent workflows. We provide end-to-end assurance across discovery to live, covering testing strategy, model and data evaluation, prompt and conversation testing, safety and responsible AI, plus security and privacy testing.

Service Description

Our approach combines automated evaluation harnesses with human-in-the-loop SME grading to provide repeatable, measurable and auditable evidence, and integrates with DevSecOps/TestOps pipelines for continuous assurance. We test accuracy, relevance, grounding, hallucination detection, harmful content, bias, fairness, prompt injection, jailbreak attempts, and data leakage across Azure OpenAI, AWS Bedrock, Google Vertex AI, and open-source models following responsible AI principles.

Service Benefits

- Validated GenAI accuracy through golden datasets and SME evaluation
- Detected hallucinations, bias, and harmful content before production deployment
- Enhanced security via prompt injection and jailbreak testing coverage
- Improved quality with automated regression testing in CI/CD pipelines
- Reduced risk through comprehensive responsible AI and safety assurance
- Better governance with auditable test evidence and transparency reports
- Accelerated deployment confidence using repeatable evaluation frameworks and harnesses

Service Features

- UK-based GenAI assurance specialists with public sector delivery experience
- Tool-agnostic testing across Azure OpenAI, Bedrock, Vertex, open-source models
- Automated evaluation harnesses for accuracy, relevance, and grounding metrics
- Human-in-the-loop SME grading for quality and relevance validation
- Responsible AI testing covering bias, fairness, toxicity, safety compliance
- Security testing for prompt injection, jailbreak, data leakage scenarios
- SC/DV eligible consultants for sensitive government GenAI assurance

Service Scope

The service covers GenAI assurance strategy, test approach design, evaluation framework implementation, automated testing, manual validation, and CI/CD integration. We specialise in chatbot testing, RAG system validation, content generation quality, summarisation accuracy, agent workflow testing, and conversation flow analysis. Work includes golden dataset creation, prompt library development, and continuous monitoring setup.

Deliverables

Typical deliverables include GenAI test strategies, evaluation frameworks, automated test harnesses, golden test datasets, prompt test libraries, bias and fairness reports, security test results,



hallucination detection reports, SME evaluation scorecards, CI/CD integration configs, evidence packs, transparency dashboards, and operational test procedures. All testing follows responsible AI and government assurance standards.

Onboarding and Offboarding

At service initiation, ITNextGen agrees scope, roles, SFIA levels, delivery approach, and reporting cadence with the buyer. Consultants are onboarded in line with buyer security and access processes. On completion, structured handover and knowledge transfer are provided, and access to buyer systems is removed in accordance with agreed offboarding procedures.

Security and Compliance

The service operates in alignment with ISO/IEC 27001 information security principles and UK government cloud security guidance. Consultants follow role-based access controls, least-privilege principles, secure change management, and incident reporting processes. Where required, consultants may be SC/DV eligible and comply with buyer-specific security requirements.

Service Constraints

Delivery is UK-based and remote-first; on-site engagement is available by agreement. Buyer access to relevant systems, environments, and data is required. Use of specific tools or platforms may require buyer-provided licences. Performance depends on timely buyer feedback, approvals, and access to subject matter experts.

Exit Management

At the end of the engagement, ITNextGen provides agreed deliverables, documentation, and knowledge transfer. All access to buyer systems is removed in line with buyer offboarding policies and contractual requirements.