



## Service Definition – itelligent-i AI (RAG and Agentic) Implementations

Service definition for: AI (RAG & Agentic) implementations

G-Cloud 15 (Lot 3 Cloud Support)

## Executive Summary (RAG-first)

itelligent-i delivers end-to-end implementation services for enterprise AI, leading with Retrieval-Augmented Generation (RAG) to provide grounded, auditable answers from your authoritative sources, and complementing with Agentic AI where autonomous, multi-step orchestration adds measurable value. Our approach is pragmatic, outcome-led, and designed for public-sector compliance from day one.

Why RAG first? RAG connects large language models (LLMs) to trusted internal knowledge and current policy/regulatory content, reducing hallucinations, improving explainability, and avoiding expensive re-training when knowledge changes frequently. We then layer Agentic AI for goal-directed workflows (plan → act → verify), governed by clear guardrails and human-in-the-loop checkpoints.

## Scope & What You Can Buy via G-Cloud

Cloud Support / AI Enablement: discovery, design, build, integration, optimisation, and managed service for RAG and agentic solutions across public cloud or hybrid.

Typical Use Cases: internal policy & guidance assistants; case-handling copilots; RFP/proposal automation; customer support deflection; compliance & risk monitoring; operational agents (triage, scheduling, data quality).

## Method & Deliverables (mapped to A–E)

### 1. Business Process Analysis & Opportunity Identification (RAG-first lens)

- Activities: Stakeholder interviews; process mapping; pain-point analysis; data/technology readiness assessment.
- Deliverables: Opportunity matrix; data readiness report; risk register; implementation options.

### 2. Sponsor Playback / Decision Gate

- Activities: Executive playback; agree success criteria and gating.
- Deliverables: Sponsor deck; options appraisal; investment case.

### 3. Rapid Proof of Concept (PoC) & Outcome Analysis

- Activities: Build RAG prototype; instrument testing framework; golden datasets.
- Deliverables: PoC demo; evaluation report; scale plan.

### 4. Full-Scale Implementation

- Activities: RAG architecture; Agentic AI orchestration; security & compliance; integration & ops.
- Deliverables: Production-ready solution; operating procedures; training materials.

#### 5. Managed Support (keeping indices & agents current)

- Activities: Content operations; agent operations; governance & reporting.
- Deliverables: Managed service; monthly reports; compliance attestations.

## Technical Approach (itelligent-i)

RAG (primary): Implement Microsoft/Azure RAG pattern; engineer for accuracy, currency, and explainability.

Agentic AI (complementary): Building on RAG, design agents that plan, act, and learn within guardrails.

## Security, Compliance & Assurance

Align to UK public-sector G-Cloud security principles and OFFICIAL data handling.

## Onboarding & Exit

Onboarding: discovery & scoping; environment readiness; approvals; sponsor playback.

Exit: data export; index snapshots; runbooks; knowledge transfer.

## Support Levels

Various support levels available as optional service.

## Pricing Model

Fixed-price workshops & PoC; time-and-materials for build; monthly recurring for support.

## Benefits & Outcomes

Grounded answers; reduced hallucination risk; faster resolution; lower TCO; scalable operations.

## Critical Success Factors

Evaluation & metrics; governance; security by design; change management; content ops; performance optimisation; scale readiness.

## Reference Architectures

RAG: ingestion → enrichment → embeddings → vector store → retrieval → prompt packaging → LLM.

Agentic: orchestrator → tools/APIs → guardrails → memory → observability.

## Pricing

Pricing is provided based on the options selected by the customer.

- Managed Services – Master data Management deployments can be incorporated into our Managed Service Provision
- All software is procured directly with vendors, itelligent-i is not a reseller.
- Support pricing varies on standard, Bronze, Silver, Gold levels. Bespoke options can be applied to suit the customers circumstances.