Waracle Ltd. - G-Cloud 15 Lot 3 Service Definition

Service: Cloud Data Platform, AI Strategy, and DataOps Consultancy

1. Service Overview and Scope

Element	Description
Service Name	Waracle Cloud Data Platform, AI Strategy, and DataOps Consultancy
Lot 3 Category	**Cloud Migration Planning** (Primary), **Set Up and Migration**, **Security Services**
Service Description	A specialist consultancy service to architect, govern, and build modern cloud data platforms. We help public sector organisations transition from legacy data warehouses to scalable Cloud Data Lakes, Mesh, or Lakehouse architectures. The service covers the full data lifecycle: strategy, engineering, governance, and AI/ML model deployment.
Target Audience	NHS Trusts, Central Government Departments, and Local Authorities seeking to modernise their data estate, implement predictive analytics, or build Secure Research Environments (TREs).
Key Supported Technologies	AWS (Glue, Redshift, Sagemaker), Azure (Synapse, Fabric, Purview), GCP (BigQuery), Databricks, Snowflake, dbt, Apache Airflow, PowerBI.

Service Scope: Inclusions and Exclusions

Area	Inclusions (What Waracle Will Do)	Exclusions (What This Service Does NOT Cover)
Strategy & Architecture	Data Maturity Assessments; Definition of Target Data Architecture (Mesh/Lakehouse); Selection of Data & AI Tooling; AI Readiness Reviews; Data Ethics Framework implementation.	Purchase of data provider subscriptions; Licensing for proprietary data platforms (e.g., Snowflake direct costs); Legal counsel on IP ownership of data.

Data Engineering (DataOps)	Design of ETL/ELT pipelines; Implementation of CI/CD for data; Data migration strategy from on-premise SQL/Oracle; Performance tuning of data warehouses.	Manual data entry or digitisation of paper records; Ongoing BAU data entry tasks.
Governance & AI	Implementation of Data Catalogs (e.g., Purview); Data Quality framework design; MLOps (Machine Learning Operations) pipeline design; Algorithmic transparency reviews.	Training of Large Language Models (LLMs) from scratch (fine-tuning is included, base model training excluded); Creation of "Deepfake" technology.

2. Outcomes and Deliverables

ID	Deliverable Title	Description and Format
D24	**Data Strategy & Roadmap**	A strategic document outlining the vision, target state, and step-by-step roadmap to achieve a data-driven organisation, aligned to CDDO recommendations.
D25	**Data Platform High-Level Design (HLD)**	Technical blueprints for the cloud data platform, detailing storage layers (Bronze/Silver/Gold), compute engines, and security boundaries.
D26	**Data Governance Framework**	A set of policies and standard operating procedures (SOPs) for data ownership, stewardship, quality standards, and catalogue management.
D27	**AI & Data Ethics Assessment**	A review of proposed AI/ML use cases against the **Government Data Ethics Framework**, identifying bias risks and transparency requirements.
D28	**Proof of Value (PoV) Report**	Summary of a pilot data project (e.g., a single use-case pipeline) to demonstrate value and technical feasibility before full rollout.

3. Compliance and Standards (Crucial for Data)

Standard	Waracle Adherence
NHS DSPT	Architectures designed to support **NHS Data Security and Protection Toolkit** compliance.
Data Ethics	All AI/ML consulting aligns with the **UK Government Data Ethics Framework** and the **AI Playbook**.
GDPR/PECR	Privacy-by-design principles applied to all data architecture (e.g., automated PII redaction, column-level security).
Open Standards	Commitment to open data standards (e.g., Parquet, Avro) to prevent vendor lock-in.