



Colibri
Digital

G-CLOUD 15 SERVICE DEFINITION

DATABRICKS LAKEHOUSE

DATA PLATFORM BUILD ON AWS

Contents

1. Service Overview	3
2. Features & Benefits	4
3. Technical Requirements & Implementation	4
4. Security & Quality Assurance	5
6. Constraints	5
7. Onboarding & Offboarding	6
Onboarding	6
Offboarding & Exit	6
Data Portability	6
Knowledge Handover	6

1. Service Overview

Supplier Name	GCI NETWORK SOLUTIONS LIMITED
Service Branding	Colibri Digital
Framework	G-Cloud 15 (RM1557.15)
Service Type	Lot 3: Cloud Support

Colibri Digital's Databricks Lakehouse Data Platform, hosted on AWS, provides a robust foundation for modern data engineering, analytics, and AI. By merging the flexibility of a data lake with the performance and governance of a data warehouse, we deliver a unified "Lakehouse" architecture directly within your AWS environment.

Our service focuses on configuring Databricks on AWS with agreed governance and security controls; assurance and compliance outcomes remain subject to buyer processes and decisions. We integrate Databricks seamlessly with native AWS services (such as S3 and IAM) to create a scalable ecosystem that eliminates data silos and reduces the total cost of ownership (TCO).

NOTE: This Service Definition covers design and build activities only; ongoing operational support and service levels are out of scope for this offering.

Key technical components of the delivery include:

Workspace Configuration	Expert setup of Databricks workspaces within your AWS VPC, including E2 architecture and PrivateLink support.
Unity Catalog & Governance	Implementation of Unity Catalog for centralised access control, auditing, and data lineage across the Lakehouse
Automated Data Ingestion	Design and build of ingestion patterns using Auto Loader and Delta Live Tables (DLT) for batch and streaming data
Infrastructure as Code (IaC)	Full platform automation using Terraform, ensuring immutable and repeatable deployments that prevent "configuration drift".
Security & Identity	Designed to support public sector security requirements, subject to buyer assurance/accreditation and agreed controls.
Operational Monitoring	Deployment of Databricks SQL warehouses leveraging built-in system tables and Query History for real-time performance and cost observability.

Core Public Sector use cases would include:

Centralised Citizen Data Records	Consolidating disparate data sources into a single "Golden Record" while maintaining a strict Unity Catalog.
Predictive Policy Modelling	Enabling data scientists to run machine learning models on a unified platform to forecast the impact of policy changes

Automated Regulatory Reporting	Modernising legacy reporting cycles through automated DLT pipelines, reducing manual effort and human error.
Secure Inter-Departmental Sharing	Facilitating the secure exchange of data between departments using governed, audited Lakehouse shares.
Real-Time Service Monitoring	Assesses the readiness of platforms to support secure data exchange between government entities. It evaluates how VPC Peering or PrivateLink can be used within a governed structure to facilitate collaborative analytics while maintaining UK data sovereignty.

2. Features & Benefits

The following table outlines the core technical features of the build and the value outcomes they support.

Technical Features	Benefits
Unified Data Architecture	Strategic Clarity: Supports data engineering, BI, and ML on a single, consistent platform for maximum ROI.
Infrastructure as Code (IaC)	Consistency & Speed: Uses Terraform to ensure environments are repeatable, version-controlled, and free from manual drift.
Unity Catalog Governance	Secure-by-Design: Centralises data access control and auditing to ensure compliance with GDPR and NCSC principles
Medallion Data Templates	Operational Excellence: Provides reusable Bronze/Silver/Gold templates, allowing teams to focus on delivering citizen-facing insights.
Comprehensive Documentation	Streamlined Accreditation: Built-in security controls simplify the process for Senior Information Risk Owners (SIROs) to approve live operations.

3. Technical Requirements & Implementation

The implementation is delivered in line with the GDS Service Standard, applying an iterative four-phase model:

The service is delivered by SC-cleared consultants, if required, using a hybrid model of remote and on-site support.

Discovery	Analysis of existing data architecture, technical debt, and sovereignty requirements to define success criteria.
Alpha	Validation of the proposed Lakehouse solution through a Proof of Concept in a sandboxed cloud environment.
Beta	Pilot migration of non-critical data workloads with limited live traffic to confirm operational readiness.
Live	Full production deployment, integration with live monitoring, and transition to steady-state management.

4. Security & Quality Assurance

As an AWS Premier Partner with the Government Consulting Competency, Colibri Digital ensures all assessments meet the highest security standards.

Secure Development:	All project staff are SC-cleared (or clearable) consultants who can be provided, subject to availability and buyer requirements.
Data Handling	Data handling and residency requirements are agreed with the buyer and implemented through the selected AWS region and agreed controls.
Standards	Our delivery workflows are to meet rigorous government and industry security benchmarks.

5. Support & Service Levels

This Service Definition covers design and build activities only; ongoing operational support and service levels are out of scope for this offering.

6. Constraints

The following constraints apply to the Databricks Lakehouse Data Platform Build on AWS, to ensure transparency and prevent disputes:

Professional Services Scope	This is a Lot 3 Cloud Support offering; it does not include ongoing AWS or Databricks consumption costs or third-party licences
Data Source Accessibility	The feasibility of automated ingestion is dependent on the availability and connectivity of legacy data systems.

7. Onboarding & Offboarding

Onboarding

Buyers are onboarded via a formal kick-off scheduled at call-off, subject to buyer availability; delivery artefacts and collaboration tooling are agreed with the buyer.

Offboarding & Exit

At completion, we hand over agreed deliverables and confirm secure return/deletion of buyer-provided data in line with agreed handling procedures.

Data Portability

All documentation is provided in open formats (PDF).

Knowledge Handover

All documentation, runbooks, and templates are handed over in open formats (PDF) to ensure operational independence.