AI-Driven Energy Optimisation & Control

Service Definition Document

G-Cloud 15 Framework

Software as a Service (SaaS) - Lot 2b

Version 1.0

January 2026

DIREK LTD

Table of Contents

1. Service Overview

1.1 What the Service Is

AI-Driven Energy Optimisation & Control is a cloud-based Software as a Service platform that transforms energy management for commercial and public sector buildings. The platform integrates existing IoT sensors and Building Management System data to deliver artificial intelligence-driven optimization of HVAC systems.

The service connects to legacy BMS infrastructure using vendor-agnostic standard protocols including BACnet/IP, Modbus TCP, and MQTT. This approach eliminates hardware replacement costs while enabling automated energy controls and environmental monitoring.

1.2 Core Capabilities

Capability Area	Description
Real-Time Analytics	Continuous monitoring and analysis of IoT sensor data, energy consumption patterns, and environmental conditions
AI Optimization	Machine learning algorithms optimize HVAC operation based on occupancy, weather, and historical performance data
Automated Controls	Direct integration with BMS enables automatic adjustment of setpoints, schedules, and control sequences
Predictive Maintenance	Anomaly detection identifies equipment faults and performance degradation before failure occurs
Compliance Reporting	Automated generation of ESG, BREEAM, ISO 50001, and NABERS compliance documentation
Agentic AI Interface	Natural language query interface allows non-technical users to extract insights and generate reports

1.3 Key Benefits

- Immediate verifiable energy savings and carbon reduction
- Zero capital expenditure through legacy system integration
- Automated compliance reduces administrative burden by hundreds of hours annually
- Data-driven maintenance prioritization based on ROI and impact analysis
- Improved indoor air quality enhances occupant health and productivity
- Unified portfolio visibility across previously siloed data sources

2. Backup, Restore, and Disaster Recovery

2.1 Backup Strategy

Automated daily snapshots capture all customer data. Backups are stored in Google Cloud Platform Storage within the London region and retained for 30 days.

Parameter	Specification
Backup Frequency	Daily (automated)
Retention Period	30 days
Storage Location	GCP Storage, London (europe-west2)
Encryption	AES-256 at rest
Restore Testing	Quarterly verification of restore procedures

2.2 Disaster Recovery Metrics

Metric	Commitment
Recovery Point Objective (RPO)	< 1 hour (maximum acceptable data loss)
Recovery Time Objective (RTO)	< 4 hours (time to restore operations)

The service architecture uses Multi-Availability Zone deployment with synchronous database replication. This configuration ensures data consistency and enables rapid failover in disaster scenarios.

3. Onboarding and Offboarding

3.1 Implementation Process

A structured 30-day Sandbox Pilot is included with each subscription. The pilot phase provides:

- Full access to all premium modules and features
- Synthetic test data for immediate system evaluation
- One-hour engineering kick-off call to configure integrations
- Technical documentation and training materials
- Email support during business hours

Pilot Exclusions: Production SLA guarantees, live PII data ingestion, and priority telephone support are not available during the pilot phase.

3.2 Data Exit and Offboarding

The service implements a vendor lock-in free approach to data ownership and portability.

Activity	Details
Data Export Formats	CSV and JSON via API (available at any time)
Export Window	30 days following service termination
Data Destruction	Cryptographic erasure 30 days after termination
Confirmation	Written certification of data destruction provided upon request

4. Service Constraints

4.1 Technical Prerequisites

Requirement	Specification
Network Connectivity	Secure outbound internet connection (HTTPS port 443, MQTT port 8883)
BMS Integration	Support for BACnet/IP, Modbus TCP, MQTT, or accessible API
Firewall Configuration	Outbound traffic permitted on standard ports

4.2 Maintenance Schedule

Planned maintenance is performed outside core UK business hours (Monday to Friday, 09:00 to 17:00 excluding public holidays). Customers receive 48 hours advance notification via email for all scheduled maintenance activities.

Emergency maintenance required for security vulnerabilities or critical system stability may be performed with reduced notice periods. Customers are notified as soon as practically possible.

4.3 Customization Limitations

The platform is delivered as a standard multi-tenant SaaS solution. Configuration options are available through the administrative interface including:

- Custom dashboards and reporting templates
- Alerting thresholds and notification rules
- User roles and permission structures
- Integration parameters and data mapping

Deep architectural modifications or bespoke feature development are not supported within the standard service model. Organizations requiring custom functionality should contact our professional services team.

5. Service Levels

5.1 Availability Commitment

Service Level Agreement: 99.9% uptime at any point (24 hours per day, 7 days per week, 365 days per year).

Availability is calculated monthly across all hours. Scheduled maintenance windows with 48 hours advance notification are excluded from availability calculations. Maintenance is performed outside core UK business hours (Monday to Friday, 09:00 to 17:00) wherever possible.

5.2 Service Credit Structure

Service credits are applied to the monthly subscription fee when availability targets are not met. Credits are calculated as follows:

Monthly Uptime	Service Credit	Application
99.0% to < 99.9%	10%	Next month's invoice
95.0% to < 99.0%	25%	Next month's invoice
< 95.0%	50%	Next month's invoice

Important: Service credits are applied automatically to the following billing cycle. Credits do not accumulate beyond one month and are not redeemable for cash.

5.3 Support Response Times

Support is delivered via email and online ticketing system. All tickets receive unique reference numbers for tracking and audit purposes.

Priority	Definition	Response Time
P1 Critical	Complete system outage affecting all users	**1 hour (24/7)**
P2 Major	Major functionality unavailable or degraded	**4 hours (business hours)**
P3 Minor	General enquiries or minor issues	**1 business day**

Business Hours: Monday to Friday, 09:00 to 17:00 (excluding UK public holidays)

Accessibility: The support ticketing system conforms to WCAG 2.2 AA standards. Users can manage ticket status and priority through the web interface.

6. Technical Requirements

6.1 Supported Browsers

Browser	Supported Versions
Google Chrome	Current version and previous major version
Mozilla Firefox	Current version and previous major version
Microsoft Edge	Current version and previous major version
Safari	Current version on macOS and iOS
Requirements	JavaScript enabled. Third-party cookies not required.

6.2 Network and Connectivity

Bandwidth: Minimum 2 Mbps per concurrent user recommended for optimal dashboard performance.

Protocol Requirements: HTTPS (port 443) and MQTT over TLS (port 8883). Firewall rules must permit outbound connections on these ports.

6.3 Authentication Methods

Method	Details
Username & Password	Standard authentication with enforced password complexity
Multi-Factor Authentication	Available for all users. Strongly recommended.
Single Sign-On (SSO)	SAML 2.0 and OpenID Connect (OIDC) protocols
Supported Identity Providers	Azure Active Directory, Okta, and others upon request

7. Hosting and Security

7.1 Infrastructure and Hosting

Component	Specification
Cloud Provider	Google Cloud Platform (GCP)
Region	London (europe-west2)
Data Residency	All data stored and processed within the United Kingdom
Architecture	Multi-Availability Zone (Multi-AZ) across 3 zones
Load Balancing	Automatic traffic distribution across availability zones

Database	Cloud SQL High Availability with synchronous replication

7.2 Architecture Diagram

System Architecture Overview:

The following describes the multi-tier architecture deployed across Google Cloud Platform:

- **User Access Layer:** End users connect via HTTPS (TLS 1.2+) from standard web browsers
- **Security Perimeter:** Google Cloud Armor Web Application Firewall (WAF) provides DDoS protection and filters malicious traffic
- **Load Balancer:** Google Cloud Load Balancing distributes incoming requests across multiple zones
- **Application Tier:** Compute instances deployed in Zone A, Zone B, and Zone C within europe-west2
- **Data Tier:** Cloud SQL Primary instance with synchronous Standby Replica for automatic failover
- **Integration Layer:** MQTT/BACnet/Modbus gateways receive data from customer BMS infrastructure

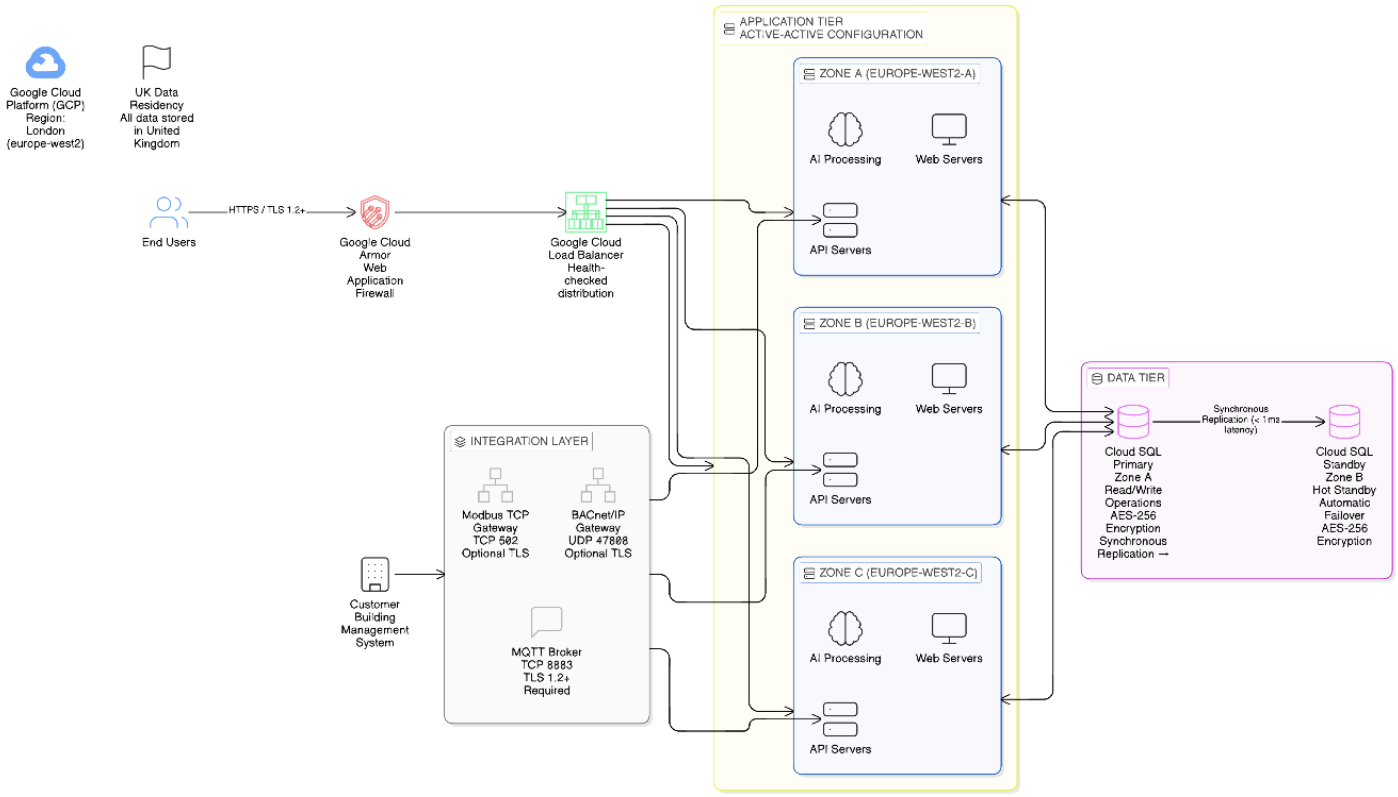

Figure 1: **High Availability Cloud Architecture (London Region)**

7.3 Data Encryption

State	Encryption Method
Data in Transit	TLS 1.2 or higher. Legacy protocols (SSL, TLS 1.0/1.1) disabled.
Data at Rest	AES-256 encryption for all databases and storage
Key Management	GCP Key Management Service with automatic key rotation
Backup Encryption	AES-256 applied to all backup snapshots

7.4 Security Standards and Compliance

ISO 27001:2022: Architecture and operational processes are aligned with ISO 27001:2022 information security management standards. Formal certification is in progress.

Secure Development: All development practices conform to recognized standards including NCSC guidance and OWASP recommendations. Self-assessed and subject to regular internal audit.

Change Management: Strict segregation between Development, Staging, and Production environments. Mandatory peer review and automated CI/CD testing before deployment. Full audit trail for all changes.

7.5 Staff Security Vetting

All personnel with access to customer data undergo background screening:

- **BS7858:2019:** Screening conforming to British Standard for security screening of individuals
- **BPSS:** Administrative staff hold Baseline Personnel Security Standard clearance
- Identity verification, employment history (3 years), right to work, and basic criminal record checks

7.6 Access Controls

Least privilege model enforced across all systems:

- Corporate identity authentication with mandatory hardware-backed MFA
- Google Identity-Aware Proxy (IAP) verifies identity and device context
- Role-Based Access Control (RBAC) via Google Cloud IAM
- Quarterly access rights review
- Time-bound 'Break Glass' emergency access with senior approval

7.7 Audit Logging

Attribute	Specification
Log Retention	Minimum 12 months for all audit logs
Log Integrity	Immutable logs prevent tampering or deletion
User Access	Real-time access via web interface and API
Events Captured	Authentication, data access, configuration changes, administrative actions

7.8 Vulnerability Management

Automated daily scanning identifies known vulnerabilities:

- Google Cloud Artifact Registry and Security Command Center scan container images
- Software Composition Analysis (SCA) detects third-party library vulnerabilities
- Remediation timelines based on CVSS severity:

Severity (CVSS)	Score Range	Patch SLA
Critical	9.0 - 10.0	**< 48 hours**
High	7.0 - 8.9	**< 7 days**
Medium	4.0 - 6.9	**< 30 days**

7.9 Incident Management

Security incidents are managed following the NCSC incident management lifecycle:

- **Detection:** Google Security Command Center provides real-time anomaly detection
- **Triage:** Immediate severity verification
- **Containment:** Isolation of affected resources to prevent spread
- **Eradication:** Vulnerability patching and threat removal
- **Recovery:** Service restoration from clean backups
- **Notification:** Customer and regulator notification within 72 hours (GDPR compliance)

Response Times: Critical/High severity incidents: 1 hour response (24/7). Medium/Low severity: 4 hours (business hours).

Post-Incident Review: Transparent PIR provided within 5 working days for all Critical/High severity events. Documents root cause, remediation, and preventative measures.