Data Classification Policy

Data Classification Policy: Orbiq

Effective Date: April 13, 2026 (rev. June 1, 2026 — cloud migration + Plaid; rev. August 16, 2026 — aggregator credentials generalised)
Document Status: Version 1.2 (Official)
Supersedes: v1.1 — retained at `archive/Data_Classification_Policy_v1.1_effective_2026-06-01.md`
Compliance Mapping: CSA CCM-DSP-04, PIPEDA Principle 7

1. Purpose

The purpose of this policy is to establish a framework for classifying data based on its sensitivity and the impact of its unauthorized disclosure. This classification ensures that Orbiq applies appropriate technical safeguards—specifically **Cloud SQL CMEK encryption**, **Postgres Row-Level Security (RLS)**, and **GCP Secret Manager**—proportionate to the risk associated with each data tier.

> **Migration note (June 2026):** Production storage is **Cloud SQL for PostgreSQL** (Canada region), not local SQLite/SQLCipher. Data-at-rest encryption is provided by Cloud SQL's CMEK/Google-managed AES-256 keys; secrets live in GCP Secret Manager. Prior references to SQLCipher and Infisical are superseded.

2. Scope

This policy applies to all data processed, stored, or transmitted by the Orbiq application, including data in the **Cloud SQL (PostgreSQL)** database, cloud-hosted logs, and secrets managed via GCP Secret Manager and third-party processors.

3. Classification Tiers

Orbiq categorizes data into four distinct tiers.

Tier 1: Public

Information that can be freely disclosed without risk to users or Orbiq.

- **Examples:** Currency exchange rates (Frankfurter API), generic merchant category lists, public marketing assets, and application version metadata.
- **Storage Requirement:** Plaintext allowed in SQLite.

Tier 2: Internal

Operations-related information that is not intended for the public but has low impact if disclosed.

- **Examples:** Non-sensitive server configurations, anonymized system performance metrics, and general UI preferences (e.g., theme settings).
- **Storage Requirement:** Plaintext allowed in SQLite; access restricted to authenticated system accounts.

Tier 3: Confidential

Sensitive information that could cause reputational or minor financial harm if disclosed.

- **Examples:** De-identified spending trends, internal architectural diagrams, and non-PII system logs.
- **Storage Requirement:** Encrypted at rest where feasible; strict access control via IAM.

Tier 4: Restricted (Primary Encryption Target)

Data requiring the highest level of protection due to its sensitive nature. Unauthorized disclosure could result in significant harm to the individual or Orbiq.

- **Examples:**
- **Personal Information (PII):** User names, email addresses, and account identifiers.
- **Financial Data:** Raw merchant strings, transaction dates, and exact currency amounts.
- **Secrets & Credentials:** API tokens (Groq, Stripe, Plaid), OAuth secrets, the aggregator token-encryption key (`PLAID_TOKEN_ENC_KEY`, which also encrypts LunchFlow keys), and database credentials.
- **Aggregator Access Credentials** — *any* credential that grants read access to a user's bank data, regardless of who issued it or who holds the vendor contract. This class is deliberately written by capability rather than by vendor name, because the previous vendor-named wording (`Plaid Connection Data`) left the LunchFlow key unclassified for the entire period it was live. Currently in scope:
- Plaid `access_token`s and item / account identifiers (Orbiq holds the vendor contract).
- **User-supplied LunchFlow REST API keys** (`lunchflow_connections.api_key_enc`) and LunchFlow account identifiers (the *user* holds the vendor contract — see RoPA §2.1). The key is the credential equivalent of a Plaid access token and is classified identically: a leak exposes the same bank data, and the fact that the user pays the vendor directly changes nothing about the blast radius.
- Any future aggregator rail. **A new integration is Tier 4 on the day the credential column is created, not on the day someone remembers to update this list.**
- **Institution-adjacent identifiers:** account numbers or masks, institution identifiers, and provider-side account IDs that can be correlated back to a named individual's bank relationship.
- **Audit Logs:** Detailed admin action logs and security-related event logs.
- **Storage Requirement:**
- **Database: Cloud SQL for PostgreSQL** with CMEK/Google-managed AES-256 at-rest encryption and **Row-Level Security (RLS)** tenant isolation. Aggregator access credentials are additionally Fernet-encrypted at the application layer before persistence, decrypted in memory only at call time, never logged, and never returned to a client. Both `plaid_items` and `lunchflow_connections` are owner-scoped RLS tables.
- **Secrets:** All API keys and environment-level secrets stored in **GCP Secret Manager** (the project secret store), fetched at runtime via Workload Identity.
- **Access:** Access is governed by the Principle of Least Privilege (PoLP) and identity-based auth (Google OAuth + IAM, MFA-enforced).

4. Data Handling & Security Controls

Tier	Encryption at Rest	Access Control	Retention Rule
Public	Optional	Open (Read-Only)	Permanent / N/A
Internal	Optional	System Admin Only	1 Year
Confidential	Recommended	Role-Based (RBAC)	1-7 Years (per RoPA)
Restricted	**Mandatory (Cloud SQL CMEK + RLS; Secret Manager)**	Least-Privilege (Workload Identity / IAM)	Defined by Retention & Deletion Policy

5. Responsibilities

- **Privacy Officer:** Responsible for maintaining this policy and conducting annual data classification audits.
- **Development Team:** Responsible for ensuring new database columns or external service integrations are classified and secured according to these tiers during the design phase.

6. Policy Enforcement

Failure to comply with this classification policy, such as storing "Restricted" data in plaintext or "Public" repositories, will trigger an immediate Incident Response assessment as defined in the Orbiq IRP.