PROJECT HELIOS DATA GOVERNANCE SPECIFICATION

Document: Helios_Data_Governance_Specification_v3.5.pdf

Status: Approved | Version 3.5 | Effective Date: October 15, 2023

Owner: Privacy Council & Head of Product

Audience: Engineering, Data Science, Product, Legal, Compliance

1.0 PURPOSE & SCOPE

This document defines all business rules, privacy constraints, and data handling requirements for Project Helios. It is the single source of truth for what data can be processed and how. All engineering implementations must map directly to clauses herein.

2.0 DATA CLASSIFICATION

2.1 Classification Levels

Level 1 (Public): Behavioral events, usage metrics

Level 2 (Internal): Aggregated customer segments

Level 3 (Confidential): Personal identifiers, contact info

Level 4 (Restricted): Financial data, health information

2.2 PII Handling Rules

Storage: Level 3/4 encrypted at rest (AES-256)

Transit: TLS 1.3 mandatory

Access: Role-based with 2FA for Level 4

Retention: 25 months behavioral data, 7 years audit trails

3.0 DATA SOURCE RULES

Source	Max Frequency	Late Data	PII Allowance
Web Analytics	Real-time	5 min	Session IDs only
Mobile App	Real-time	10 min	Device ID + consent
Partner APIs	15-min batches	1 hour	Contractual fields
Support Tickets	5-min batches	2 hours	Contact info (masked)

4.0 PROFILE MERGING LOGIC

4.1 Matching Thresholds

Match Type	Confidence	Action
Email → User ID	99%	Auto merge
Device ID → Anonymous	85%	Auto merge
Household Inference	75%	Flag for review
Cross-Device	70%	Manual approval only

4.2 Conflict Resolution Order

Direct Customer Input (Support, account updates)

Transactional Systems (Payment data)

Behavioral Data (App/web usage)

Inferred Data (ML predictions)

4.3 Field Precedence

Field	Primary Source	Secondary	Merge Strategy
Email	Account System	Partner API	Use Primary
Phone	Account System	Support	Most recent
Preferences	Behavioral	Survey	70/30 weighted
Address	Transaction	Partner	Manual verify

5.0 DATA QUALITY SLAS

5.1 Processing Latency

Ingestion: < 1000ms to Kafka (P95)

Profile Update: < 500ms (P95)

Allowed Lateness: 10-minute watermark

5.2 Quality Metrics

Completeness: > 98% mandatory fields

Accuracy: > 95% cross-verified

Timeliness: 99% within SLA windows

Validity: 100% schema compliance (Level 3/4)

6.0 PRIVACY CONTROLS

6.1 Right to Erasure

Anonymize Level 3/4 within 24 hours

Retain aggregated behavioral data

Maintain 7-year audit trail

Propagate to downstream in 48 hours

6.2 Consent Management

Explicit: Marketing, third-party sharing

Implied: Service improvement, security

Withdrawal: < 1 hour processing

Audit: 7-year immutable log

6.3 Data Sovereignty

EU Data: Frankfurt/Dublin only

US Data: Virginia/Oregon only

APAC Data: Singapore/Sydney only

Cross-Region: Legal approval + extra encryption

7.0 MONITORING & AUDIT

7.1 Mandatory Logs Per Operation

Timestamp (UTC)

Operation type (CRUD/Merge)

User/Service principal

Fields accessed/modified

Rule reference (e.g., "Section 4.2")

Source system ID

7.2 Alert Thresholds

Critical: Match confidence < Table thresholds

High: Data quality < 90% for 15+ minutes

Medium: Latency > SLA for 5+ minutes

Low: Schema failures > 1% traffic

8.0 CHANGE MANAGEMENT

8.1 Amendment Process

Proposal with business justification

Privacy Council review (2 days)

Legal compliance check

Engineering impact assessment

Version update + documentation

7-day notice before implementation

8.2 Emergency Changes

Security vulnerabilities only:

CTO + Privacy Head joint approval

Immediate implementation

Retroactive docs within 24 hours

Stakeholder notice within 48 hours

9.0 TRAINING REQUIREMENTS

Engineers must certify understanding of:

Sections 4.0 (Profile Merging)

Sections 6.0 (Privacy Controls)

Appendix B (Classification Matrix)

Related Documents:

Architecture: Helios_Platform_Architecture_v4.0.pptx

Incident Response Playbook

Data Breach Notification Procedure

APPROVAL SIGNATURES

Head of Privacy: ____________________ Date: _________

CTO: ___________ Date: _________

Head of Product: ____________________ Date: _________