

ODA3 INSTITUTE
TECHNICAL REFERENCE

UAIF Security and Privacy Handling Guide

Handling Sensitive Data Within UAIF Incident Records

TCR-UAIF-023 · Version 1.0
Status: Published · Adoption-Phase Reference

Governs: uaif-core-v1.0.schema.json field-level data handling
Framework: Unified AI Incident Framework (UAIF™) v1.0

© 2026 ODA3 Pvt Ltd · Published as ODA3 Institute
oda3.org · CONTACT_AT_ODA3_DOT_ORG

Document Control and Metadata

Document identifier	TCR-UAIF-023
Doc Type	Technical Reference — Security and Privacy Handling Guide
Title	UAIF Security and Privacy Handling Guide
Version	1.0
Publication date	14 July 2026
Review cycle	Annual, or on any change to applicable data-protection law
Publisher	ODA3 Institute
Legal entity	ODA3 Pvt Ltd
Classification	Open — GAISSE Ecosystem source-available material
Framework tag	UAIF™ v1.0
Regulatory crosswalk	GDPR (EU 2016/679); India DPDP Act 2023; sector breach-notification regimes referenced in regulatory_obligations
Evidence tier	T1–T2 for legal citations; remainder is schema-derived operational guidance
Incident provenance	Not applicable
Audience	Implementers, SOC/GRC engineers, privacy and security reviewers handling UAIF records
Licence	Subject to the GAISSE Ecosystem Licence (GEL) v1.0
Contact	CONTACT_AT_ODA3_DOT_ORG · oda3.org

SCOPE NOTICE

A UAIF incident record is itself a security- and privacy-sensitive artefact: it can describe an active or recent AI security failure, and its fields can end up carrying personal, confidential, or classified information if implementers are not deliberate about what they populate. This guide addresses handling of the record, not the design of the AI system that produced the incident.

1. Purpose and Scope

This guide describes how to handle sensitive data that may appear in, or be implied by, a UAIF Core Profile record: incident narratives, exploit descriptions, x_extensions payloads, and metadata that could re-identify a person, system, or organisation. It applies to any implementer, integrator, or analyst who creates, stores, transmits, or shares UAIF records.

It does not replace an organisation's own data-protection impact assessment, security classification policy, or incident-response plan; it addresses UAIF-specific handling questions those broader policies do not cover.

2. Fields That Commonly Carry Sensitive Content

Field	Sensitivity risk	Handling guidance
incident_summary / exploit_path_description	Free-text fields (20–4000 characters) can easily carry names, credentials, IP addresses, internal hostnames, or excerpts of exfiltrated data if the analyst pastes raw evidence.	Redact or pseudonymise before populating. Reference an internal evidence store by identifier rather than pasting raw logs or payloads.
x_extensions	An open object with no schema-level content restriction — sector calibration extensions or ad hoc fields can carry anything an implementer puts there.	Treat every x_extensions key as unreviewed until your own data classification confirms otherwise. Do not assume schema validation provides any privacy or security screening — it does not.
severity_rationale.assumptions / precision_notes	Free-text analyst commentary can inadvertently include speculation about specific individuals (e.g. suspected insider, named victim).	Keep to technical/process rationale; avoid naming individuals unless the record's access controls specifically permit it.
linked_vulnerability_id / mcp_endpoint_ids / oauth_pivot_chain / agent_escalation_chain	Identifiers here can reveal internal system topology, endpoint names, or credential-adjacent infrastructure detail.	Use internal reference identifiers rather than literal endpoint URLs, hostnames, or credential-adjacent strings.
jurisdiction_country_code / sector_code / affected_counts (where sector extensions add them)	Individually low-risk, but in combination with a public disclosure timeline can support re-identification of a specific organisation or incident.	Apply your own disclosure-timing and aggregation review before any external sharing, independent of schema validation passing.

3. Do Not Submit — Hard Boundaries

Consistent with the GAISSE Ecosystem's contribution and licensing instruments (GEL v1.0 Section 8; UAIF-CLA-002 Section 10), the following SHALL NOT be placed into any UAIF record intended for sharing outside a strictly controlled, access-limited environment:

- personal data, personally identifiable information, or protected health information;
- payment-card data, authentication credentials, access tokens, or security keys;
- unredacted third-party confidential information;
- classified or export-controlled material; or
- biometric data, except as an aggregate boolean flag (e.g. biometric_system_involved) rather than the biometric data itself.

Where such data is unavoidably relevant to an incident (for example, a privacy breach whose harm depends on what data was exposed), record the fact and category of exposure using the schema's

boolean and enum fields (sensitive_data, harm categories, x_extensions flags) rather than the underlying data itself.

4. Access Control and Sharing

- Apply the principle of least privilege to who can read a UAIF record's free-text fields versus its structured/enum fields — the structured fields are far lower-risk to share broadly (e.g. for aggregate reporting) than the narrative fields.
- When sharing records externally (e.g. for sector benchmarking or regulatory reporting), strip or generalise incident_summary and exploit_path_description rather than sharing them verbatim, unless the specific recipient and legal basis justify full disclosure.
- Records flagged regulatory_trigger = true, or carrying vulnerable_population / critical_sector / sensitive_data = true, warrant expedited internal security and privacy review before any sharing decision, independent of the severity score.
- Maintain an audit trail of who accessed or exported a record, separate from the UAIF record's own fields — UAIF does not define an access-log schema.

5. Retention and Deletion

UAIF does not define a retention period; retention is governed by the implementer's own data-protection obligations and incident-response policy. Where a jurisdiction's law requires deletion or the right to erasure applies, implementers should be aware that deduplication_hash and incident_uuid are structural identifiers, not personal data themselves, but the record's narrative fields may need redaction or deletion independent of the structural record's retention for audit-trail purposes.

6. Regulatory Cross-References

Instrument	Relevance
GDPR (EU) 2016/679, Articles 5, 32–34	Data minimisation, security-of-processing, and breach-notification obligations apply to any UAIF record containing personal data, independent of the record's own regulatory_obligations field.
India Digital Personal Data Protection Act, 2023	Applies to processing of personal data referenced in or by a UAIF record where the Act's territorial scope is engaged.
Sector-specific breach regimes (e.g. GDPR Art. 33, NIS2 Art. 23, DORA Art. 19)	The schema's regulatory_obligations enum flags these routes for screening; it does not perform the legal assessment itself — see the applicable Sector Calibration Package's routing matrix.

7. Notably Absent

- This guide does not provide a data-loss-prevention or redaction tool; redaction remains a manual or tooling responsibility outside UAIF.

- It does not define an access-control or audit-log schema — those remain the implementer's own infrastructure responsibility.
- It does not constitute legal advice on any jurisdiction's data-protection obligations; consult qualified counsel for a specific fact pattern.
- It does not cover security hardening of the systems that store or transmit UAIF records (encryption at rest/in transit, key management) — that is general information-security practice, not UAIF-specific.