

Strategic Proposal Outline: OKF-CM

Open Knowledge Format for Configuration Management (OKF-CM):
Deterministic Hashing, Cryptographic Attestation, and Directed Graph
Semantics for AI-Governed Engineering Baselines

Executive Summary & Strategic Rationale

Pillar 1: The Essential Role of CM in AI System Training & Context

AI models and autonomous agents do not suffer from a lack of data; they suffer from a lack of *governed context*. Prompting or fine-tuning AI tools on raw, un-curated file trees leads to structural hallucinations, context drift, and non-deterministic behavior.

Configuration Management (CM) provides the exact structural discipline required to transform raw enterprise documentation into a trusted knowledge engine. **OKF-CM** establishes the formal boundary conditions, metadata wrappers, and ownership chains needed to feed AI models safely.

Pillar 2: Strict Version Authority & Lifecycle Isolation

An AI agent operating in an engineering or defense environment **must only read, train, and execute against active, verified Configuration Item (CI) versions**.

Allowing an AI agent to index deprecated specs, preliminary drafts, or tampered artifacts risks catastrophic execution errors. **OKF-CM** enforces machine-parsable status enums (**approved**, **deprecated**, **superseded**), strict schema rules, and Merkle-tree cryptographic signatures (\$KOH\$) so that AI parsers immediately reject invalid or outdated baselines before ingestion.

Pillar 3: Interoperability between the CM "Vault of Knowledge" and Google's OKF Standard

Google introduced the Open Knowledge Format (OKF) as an open standard for AI agents to read and traverse structured Markdown knowledge bases.

By defining **OKF-CM**, we bridge the aerospace and defense CM "Vault of Knowledge" directly into the native language of AI agents. Google provides the general-purpose knowledge representation framework; **OKF-CM** adds the zero-trust cryptographic rigor, multi-tier Merkle trees, and explicit graph semantics (**depends_on**, **satisfies**, **supersedes**) required for mission-critical digital engineering.

Technical Architecture Mapping

- **Google Open Knowledge Format:** Base Specification for Human/AI Markdown Knowledge
 - **OKF-CM Profile:** Aerospace Configuration Management Specification
 - **Pillar 1:** Manifest Schema & AI Ingestion Boundary (Strict JSON Schema)
 - **Pillar 2:** Assurance & Deterministic State (3-Tier Merkle Trees & Signatures)
 - **Pillar 3:** ORDER Graph & Traversal Semantics (Directed Multi-Graph Edges)

Detailed Specification Refinements

1. OKF-CM-001: Manifest Schema & Ingestion Safety (Pillar 1)

- **Goal:** Ensure raw directory trees are safe for automated AI consumption.
- **Mechanism:** Every Knowledge Object directory must contain an `.okf.md` manifest enforcing UTF-8 encoding, standardized YAML frontmatter, and a Markdown body.
- **AI Security Boundary:** Implements strict directory traversal checks (`OKF_ERR_PATH_TRAVERSAL`) to ensure an AI parser or tool cannot escape the governed directory root via malicious relative path escaping (`..`).

2. OKF-CM-002: Assurance & Deterministic State / ADS (Pillar 2)

- **Goal:** Guarantee absolute version authority and detect state drift prior to AI ingestion.
- **Mechanism:** Computes a single 32-byte **Knowledge Object Hash (\$KOH\$)** using a 3-tier Merkle Tree:
 - **Tier 1 (\$H_{\text{artifacts}}\$):** Lexicographically sorted byte hashes of all payload files.
 - **Tier 2 (\$H_{\text{manifest}}\$):** Normalized line endings of `.okf.md` excluding the signature.
 - **Tier 3 (\$KOH\$):**
 - $$$KOH = \text{SHA256}(H_{\text{manifest}} \mathbin{\text{Vert}} H_{\text{artifacts}})$$$
- **Lifecycle Isolation:** If `status` is not `approved`, or if the calculated `$KOH$` does not match the cryptographic signature, AI ingestion agents **must fail fast** (`OKF_ERR_KOH_MISMATCH`).

3. OKF-CM-003: ORDER Graph & Traversal Semantics (Pillar 3)

- **Goal:** Allow AI agents to perform complex impact analysis and dependency tracing over local filesystems without needing proprietary or cloud-hosted database indexers.
- **Mechanism:** Defines a Directed Multi-Graph $G=(V,E)$ with strict inverse mapping:
 - `depends_on` $\rightarrow$ `required_by`
 - `satisfies` $\rightarrow$ `satisfied_by`
 - `supersedes` $\rightarrow$ `superseded_by`
- **Agent Capabilities:** AI agents can traverse local `.okf.md` relationship blocks outbound or inbound to calculate the exact blast radius of a proposed engineering change.

Committee Value Proposition Summary

- **For SAE S-19 (Digital Engineering):** Delivers a concrete, file-system-native specification that brings machine-readable CM directly into modern digital thread and AI agent workflows.
- **For SAE G-33 (Configuration Management):** Reasserts the vital necessity of CM in the modern technology stack, proving that AI safety and accuracy rely entirely on classic CM principles of baseline integrity and version authority.