Project GAINS

(Generative AI for Network Sustainability)

Deliverable D3.3

Technical Report on Design and Implementation of Knowledge Manager

Giordano, A. – Basile, L. – De Giacomo A. – Haider, B. – Fruncillo, D.

2026-01

Work Package: WP3 Design, Prototype Implementation and Evaluation of Agents

Status: Final | Version: 1.0

Classification

Abstract

This work describes the design and prototypical implementation of the Knowledge Manager, a knowledge-plane component oriented toward operations for telco and cloud environments, with the objective of transforming documentation, policies, and technical artifacts into a system that is queryable, applicable, and verifiable. The Knowledge Manager is conceived as a containerized service, reproducible in Docker, capable of operating with an external OpenAI-compatible endpoint for embeddings and generation, maintaining decoupling from the provider and control over the information perimeter effectively transmitted externally. The design is centered on three non-negotiable requirements: contextual applicability, grounding with verifiable citations, and security against contamination and exfiltration. The solution adopts a layered architecture: deterministic ingest pipeline, knowledge store with hybrid search and structured filters, and agentic runtime for retrieval and synthesis. The ingest is implemented in Langflow through versioned flows, exportable and importable, which perform parsing, normalization, deterministic chunking, canonicalized hashing, mandatory redaction for snapshots, and anti-leakage controls with fail-closed rejection. Indexing uses Weaviate with BYOV configuration, in which vectors are computed by the application and associated with persisted objects, in order to preserve portability and governance of the embeddings lifecycle. Retrieval combines keyword search and vector search through hybrid search, with selectable fusion algorithms and inspectable scoring metadata, and applies binding filters based on flattened applicability properties such as vendor, device role, service, and version constraints, reducing the probability of semantic drift in operational contexts. Agentic behavior is orchestrated through Elysia, adopting a decision tree that makes explicit the transitions between request classification, gating, multi-source retrieval, deterministic re-ranking, construction of the evidence pack, grounded synthesis, and final validation. The operational mode imposes strict constraints: no prescriptive claim is emitted without at least one valid reference, and conflicts between sources are resolved with a policy-first approach, degrading the response to non-operational when evidence is insufficient. To support auditability, the system produces an Explainability Pack that can be persisted and exported, which includes claim-refs mapping, retrieval parameters, artifact hashes, and quality indicators such as citation coverage and unsupported claim rate, enabling repeatable verification and diagnosis of regressions.

Consistent API contracts are defined for ingest, cited or structured query, service blueprint, snippet discovery, and explain pack, with a common envelope, error model, RBAC, and export controls. The prototypical implementation includes a reference repository, full docker-compose, automatic bootstrap of schema and flows, secret management via Docker secrets, and reproducible smoke tests. Evaluation is structured on retrieval and grounding metrics, including precision and stability of top-k, applicability violation rate, citation coverage, and absence of unsupported claims in operational mode, with a regression pipeline on a locked dataset. Security is treated as an end-to-end requirement: minimization of payloads toward the LLM, defenses against direct and indirect prompt injection through evidence-pack isolation, export control policies, rate limiting, and fail-closed behavior, in alignment with the main risk taxonomies for LLM-integrated applications. Overall, the Knowledge Manager provides a solid implementation foundation for integrating technical knowledge and operational governance in an agentic context, with industrial reproducibility, traceability, and verifiable quality controls.

Table of Contents