

Enterprise Knowledge Graph Architecture

A Comprehensive Enterprise Architecture Approach for Business Knowledge, Semantic Integration, Graph Analytics, AI Grounding, RAG, and Decision Intelligence

Sanjeeve Kumar Gajadi*

*(HCL Technologies, Bengaluru, Karnataka, India)

Email: sanjeevekumar.g@hcltech.com

Abstract:

Enterprise information is distributed across applications, data platforms, documents, metadata repositories, process models, and architecture tools, making relationship-centric analysis difficult. This paper proposes an Enterprise Knowledge Graph Architecture that connects business entities, semantics, metadata, lineage, processes, policies, technologies, risks, and events into a governed relationship layer. The framework integrates ontology management, entity resolution, graph storage, data ingestion, provenance, security, query services, graph analytics, SAP Business Data Cloud, SAP Datasphere, SAP HANA Cloud, SAP LeanIX, SAP Signavio, Retrieval-Augmented Generation, and Artificial Intelligence. It explains how Knowledge Graphs support enterprise search, architecture impact analysis, data discovery, decision intelligence, AI agents, and graph-enhanced RAG while preserving ownership, access control, semantic quality, and lifecycle governance. The proposed architecture enables organizations to transform fragmented information into reusable Business Knowledge and improve contextual intelligence across enterprise domains.

Keywords — Knowledge Graph, Enterprise Architecture, Business Knowledge, Ontology, Semantic Layer, SAP BTP, SAP Business Data Cloud, SAP Datasphere, SAP HANA Cloud, Graph Analytics, RAG, Artificial Intelligence, Enterprise Metadata, Decision Intelligence

I. INTRODUCTION

Enterprise Knowledge Graph Architecture provides a structured way to connect business entities, relationships, semantics, policies, processes, events, and metadata across complex digital landscapes. Traditional data platforms organize information primarily into tables, files, and application-specific schemas. Knowledge Graphs add an explicit relationship and semantic layer that enables enterprise information to be understood in context.

The architecture is particularly valuable for Artificial Intelligence because Generative AI, RAG, recommendation systems, fraud analytics, enterprise search, and decision intelligence all benefit from trusted relationships and business meaning.

II. BUSINESS PROBLEM

Large enterprises typically operate hundreds or thousands of applications with duplicated definitions of customers, products, assets, suppliers, employees, contracts, risks, and processes. Information is often technically integrated but semantically fragmented.

This fragmentation makes it difficult to answer relationship-rich questions such as which suppliers support a critical product, which applications depend on a vulnerable technology, which contracts are affected by a regulatory change, or how a business capability connects to data, processes, and systems. Knowledge

Graphs provide an architecture for resolving these relationships explicitly.

III. ARCHITECTURE OBJECTIVES

The Enterprise Knowledge Graph Architecture aims to create a governed semantic layer that supports entity resolution, relationship intelligence, enterprise search, AI grounding, metadata discovery, lineage, impact analysis, and decision support.

The architecture should remain interoperable with data warehouses, lakehouses, APIs, event platforms, analytics, and operational systems. A Knowledge Graph complements rather than replaces these platforms.

IV. SEMANTIC MODEL AND ONTOLOGY

The ontology defines the concepts and relationships that give the graph business meaning. Typical enterprise classes include Customer, Product, Supplier, Asset, Employee, Application, Capability, Process, Contract, Policy, Risk, Location, Data Product, and Technology.

Ontology governance should balance enterprise standardization with domain flexibility. Core concepts and identifiers can be standardized centrally while domain extensions evolve under controlled governance. Excessive central modeling can become slow, while uncontrolled local schemas recreate fragmentation.

V. ENTITY RESOLUTION

Entity resolution links records that refer to the same real-world object across different systems. Matching may use deterministic identifiers, master data, probabilistic rules, fuzzy matching, and machine learning. Confidence scores and provenance should be retained so uncertain matches can be reviewed.

A persistent enterprise identifier allows relationships from multiple systems to converge around one governed entity while preserving source-specific attributes and lineage.

VI. DATA INGESTION AND INTEGRATION

Graph ingestion pipelines can consume structured tables, APIs, events, documents, metadata catalogs, process repositories, and external knowledge sources. SAP Integration Suite, data pipelines, APIs, and event services can synchronize relevant entities and relationships into the graph.

Ingestion should preserve source, timestamp, transformation logic, security classification, and ownership. Incremental change capture is important because stale relationships can produce incorrect AI or business decisions.

VII. GRAPH STORAGE AND QUERY

Enterprise graph platforms may use property-graph or RDF-based representations depending on requirements. Graph queries support multi-hop relationship traversal, path finding, pattern detection, neighborhood analysis, and semantic reasoning.

SAP HANA Cloud graph capabilities or other enterprise graph technologies can support relationship-centric workloads. Query APIs should abstract implementation details where reusable enterprise services are needed.

VIII. METADATA AND LINEAGE

Knowledge Graphs can unify technical metadata with business context. Data products, tables, reports, models, applications, owners, policies, and business capabilities can be linked to show impact and lineage across the enterprise.

This is valuable for architecture governance, compliance, data discovery, and change management. A technology change can be traced to dependent applications, processes, business capabilities, and business owners before implementation.

IX. BUSINESS KNOWLEDGE AND SAP ARCHITECTURE

SAP Business Data Cloud and SAP Datasphere can provide governed semantic data products, while SAP LeanIX contributes application and technology portfolio context and SAP Signavio contributes business process context. A Knowledge Graph can connect these perspectives into a relationship-rich Business Knowledge layer.

This combined context supports enterprise architecture decisions, modernization planning, process transformation, AI grounding, and executive analysis.

X. KNOWLEDGE GRAPH FOR RAG

Graph-enhanced RAG uses entities and relationships to improve retrieval beyond text similarity. The system can identify relevant nodes, traverse relationships, filter by business context, and combine graph evidence with document retrieval.

This is particularly useful for questions requiring multi-hop reasoning, such as identifying which customers may be affected by a supplier issue or which business processes depend on a technology scheduled for retirement.

XI. KNOWLEDGE GRAPH FOR AI AGENTS

AI agents can use a Knowledge Graph as a governed source of enterprise context. The graph helps agents identify allowed entities, dependencies, ownership, policies, and relationships before selecting actions or tools.

Agent access should be authorized at entity and relationship level where necessary, and graph queries should be logged for audit. Sensitive relationships require the same privacy and security discipline as source systems.

XII. SECURITY AND ACCESS CONTROL

Knowledge Graphs may expose relationships that are more sensitive than individual records. Security must therefore protect nodes, edges, attributes, query APIs, and derived inferences. Identity-based access, encryption, tenant isolation, network security, and audit logging are foundational controls.

The architecture should prevent inference attacks in which a user reconstructs restricted information through permitted relationships. Sensitive domains may require subgraph filtering, query policies, masking, or separate graph partitions.

XIII. GOVERNANCE AND STEWARDSHIP

Graph governance defines ontology ownership, data stewardship, quality rules, source authority, entity resolution policy, change management, and retirement. Every major domain should have accountable business and technical stewards.

Governance should also define how conflicting facts are handled. The graph may preserve multiple source assertions together with provenance rather than forcing premature reconciliation when business truth is context-dependent.

XIV. QUALITY MANAGEMENT

Graph quality includes entity accuracy, relationship accuracy, completeness, consistency, timeliness, provenance, and semantic validity. Automated rules can identify orphan nodes, impossible relationships,

duplicate entities, stale links, and missing mandatory properties.

Quality metrics should be visible to graph consumers so AI and analytics services can account for uncertainty. Critical relationships may require higher verification standards than exploratory knowledge links.

XV. OBSERVABILITY AND PERFORMANCE

Graph observability should monitor ingestion latency, query performance, graph growth, failed transformations, authorization errors, quality violations, and usage patterns. High-degree nodes or complex traversals can create performance hotspots that require indexing, partitioning, caching, or query optimization.

Operational monitoring should also detect unexpected changes in critical relationship patterns, which may indicate source errors or security issues.

XVI. IMPLEMENTATION METHODOLOGY

Implementation should begin with a high-value relationship-centric use case rather than an attempt to model the entire enterprise. Teams define the business questions, domain entities, authoritative sources, ontology, integration approach, security model, quality metrics, and success criteria.

After a successful pilot, common ontology components, ingestion patterns, identifiers, graph APIs, and governance processes can be standardized for reuse across additional domains.

XVII. INDUSTRY USE CASES

In manufacturing, Knowledge Graphs can connect equipment, components, suppliers, maintenance records, and engineering documentation. In financial services, they can connect customers, accounts, transactions, counterparties, and risk relationships. In life sciences, they can connect research entities, products, studies, regulations, and evidence.

In enterprise architecture, a graph can connect business capabilities, applications, technologies, interfaces, data products, risks, and transformation roadmaps to support impact analysis and modernization decisions.

XVIII. FUTURE TRENDS

Enterprise Knowledge Graphs will increasingly converge with vector search, Generative AI, semantic business data, digital twins, event streams, and agentic AI. Hybrid graph-vector retrieval will enable AI systems to combine semantic similarity with explicit business relationships.

Ontology generation and graph maintenance will also become more AI-assisted, but human governance will remain necessary for business meaning, ownership, and high-impact relationships.

XIX. CONCLUSION

Enterprise Knowledge Graph Architecture creates a semantic and relationship layer that connects fragmented enterprise information into usable Business Knowledge. It strengthens enterprise search, architecture governance, data discovery, decision intelligence, RAG, and AI agents.

By combining governed ontologies, entity resolution, provenance, security, graph analytics, and lifecycle stewardship, organizations can create a reusable knowledge foundation that improves contextual intelligence across the enterprise.

REFERENCES

- [1] ISO/IEC 42001:2023, Information technology — Artificial intelligence — Management system.
- [2] National Institute of Standards and Technology, Artificial Intelligence Risk Management Framework (AI RMF 1.0), 2023.
- [3] The Open Group, TOGAF Standard, 10th Edition, 2022.
- [4] ISO/IEC 27001:2022, Information security, cybersecurity and privacy protection — Information security management systems.
- [5] ISO 31000:2018, Risk management — Guidelines.
- [6] DAMA International, DAMA-DMBOK2: Data Management Body of Knowledge, 2017.
- [7] W3C, RDF 1.1 Concepts and Abstract Syntax.
- [8] W3C, SPARQL 1.1 Query Language.
- [9] W3C, Web Ontology Language (OWL) 2 Overview.
- [10] SAP SE, SAP Business Technology Platform Documentation.
- [11] SAP SE, SAP Business Data Cloud and SAP Datasphere Documentation.
- [12] SAP SE, SAP HANA Cloud Documentation.
- [13] SAP SE, SAP Analytics Cloud Documentation.
- [14] SAP SE, SAP Integration Suite Documentation.
- [15] SAP SE, SAP AI Core and SAP AI Launchpad Documentation.
- [16] SAP SE, SAP LeanIX Enterprise Architecture Documentation.
- [17] SAP SE, SAP Signavio Process Transformation Suite Documentation.
- [18] OECD, Recommendation of the Council on Artificial Intelligence.
- [19] IEEE, Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems.
- [20] Cloud Native Computing Foundation, Cloud Native Landscape and Kubernetes Documentation.

AUTHOR BIOGRAPHY

Sanjeeve Kumar Gajadi is an SAP Enterprise Architect and Lead Consultant with extensive experience in enterprise architecture, SAP Business Technology Platform, data and analytics, integration, cloud architecture, security, governance, and AI-enabled transformation. His professional interests include enterprise knowledge graphs, semantic

architecture, SAP BTP, Business Data Cloud, RAG, AI grounding, and decision intelligence.

ACKNOWLEDGMENT

The author acknowledges the contributions of enterprise architecture communities, international standards bodies, research communities, and SAP documentation teams whose published guidance informed the concepts synthesized in this paper.