AI Engineering
https://www.sciltp.com/journals/aieng

Article

A Five-Layer Reference Architecture for First-Party Enterprise Knowledge Graphs with GraphRAG Integration and Governance Controls

Asheesh Pandey

Independent Researcher, Jersey City, NJ 07306, USA; asheesh.pandey@outlook.com

How To Cite: Pandey, A. A Five-Layer Reference Architecture for First-Party Enterprise Knowledge Graphs with GraphRAG Integration and Governance Controls. *AI Engineering* **2026**, *2*(2), 14. https://doi.org/10.53941/aieng.2026.100014

Received: 4 May 2026
Revised: 26 July 2026
Accepted: 6 August 2026
Published: 14 August 2026

Abstract: Organizations produce significant amounts of first-party data as part of their day-to-day operations in manufacturing, clinical, and customer experience domains, which are not accessible to generic large language models. The paper discusses the five-layer approach to using this information asset: (1) data ingestion and quality assurance, (2) domain ontology engineering, (3) knowledge graph engineering and population, (4) GraphRAG-enabled AI augmentation, and (5) downstream application enablement. The paper focuses on the design considerations, implementation tactics, and lessons learned from real-world applications in manufacturing, healthcare, and professional networks rather than presenting original research results. The implementation results show up to 70–80% query time reduction and close to 85% fewer hallucinations on average for GraphRAG over the standard RAG for most of the use cases analyzed. The main barriers to adoption are the substantial manual effort involved in ontology engineering, 73–94% entity resolution accuracy across different industries, and 3–5× higher computational costs for GraphRAG compared to RAG. The framework gives practitioners and researchers a reference architecture for designing, evaluating, and governing enterprise knowledge graph deployments built on proprietary organizational data.

Keywords: knowledge graphs; first-party data; GraphRAG; enterprise AI; semantic ontology; data governance; LLM integration; entity resolution; information retrieval; privacy compliance

1. Introduction

Enterprises operate at scale and generate first-party data, records of their own activity that constitute unique knowledge assets. Knowledge workers in manufacturing waste up to 20% of their working time searching for technical information scattered across various sources [1], while engineers routinely extract maintenance and design-related data from dozens of repositories [2], a pattern that makes AI-supported decision-making harder than it should be.

The relational tables, document stores, and flat files comprising many enterprise databases capture facts, values, and records but lack connections that encode information [3]. At the same time, large language models (LLMs) introduce new complications: question-answering hallucination, access limitations to proprietary data, and insufficient domain expertise reduce the chances of obtaining value from these models [4]. According to OpenAI's survey of more than 9000 enterprise workers conducted in December 2025, access to high-quality data is the primary barrier to realizing the value of foundation models in practical use cases [5]. Connecting the facts stored in enterprise data with relational logic is a critical challenge for enterprise AI. McKinsey's 2024 survey of global AI adoption in enterprises revealed that 65% of respondents used generative AI at least once a week in their

work, compared to only 35% in 2023, and the ability to operate domain-specific knowledge graphs representing the organization's data has become a strategic priority [6].

Knowledge graphs capture information as statements about relationships between real-world entities connected with typed links. The approach provides an opportunity to ground language models' reasoning processes in the explicit information captured by the knowledge graph rather than implicit vector-space relationships. GraphRAG approach proposed by Edge et al. [7] demonstrated that using enhanced graph-based retrieval leads to dramatic gains in both the factual accuracy and interpretability of generated responses compared to standard vector-space approaches. Implementation of graph-augmented approaches in production requires a reference architecture that incorporates best practices for data capture, graph construction, and application of graph-enhanced LLMs for question-answering in specific domains. To the best of our knowledge, there is a lack of literature describing such a reference architecture for first-party enterprise data.

We describe and evaluate a five-layer reference architecture for graph-driven enterprise AI applications using first-party data. Our contribution can be summarized in systematizing the knowledge graph engineering and enterprise data management practices into a coherent set of recommendations for building domain-specific ontologies, constructing knowledge graphs at scale, and utilizing GraphRAG approaches for downstream application enablement. We analyze the application of this architecture across three different problem domains: manufacturing, healthcare, and professional networks. The key implementation challenges identified in these domains offer generalizable insights and solutions applicable to a variety of verticals and use cases.

Prior literature partially addresses the issues outlined in this work. In particular, the existing works only partially address the application of knowledge graph construction practices, enterprise data management approaches, and large language models to solve a set of related issues in a production environment. Edge et al. [7] describe the GraphRAG retrieval mechanism in great detail; however, they do not elaborate on the application of GraphRAG to enterprise data management. Hogan et al. [3] provide an in-depth overview of the knowledge graph research field; however, the integration of language models is not their primary focus. ISO/IEC 25012 [8] provides an extensive description of data quality dimensions, but it does not elaborate on the particularities of applying these dimensions within a knowledge graph framework. Moreover, none of the works jointly address the issues of knowledge graph construction, enterprise data management, and large language models to solve a set of related issues.

1.1. Research Objectives

This study has four objectives:

(1) Propose an enterprise context-aware knowledge graph architecture based on first-party data
(2) Analyze property graph integration patterns with large language models and evaluate GraphRAG retrieval approaches
(3) Assess quality, governance, and privacy aspects of first-party data for production AI
(4) Identify critical technical and organizational barriers to enterprise knowledge graph adoption and review potential solutions.

1.2. Principal Contributions

This work advances the field of enterprise AI and knowledge graphs in three ways:

(1) Presenting a five-layer reference architecture for enterprise context graph systems that standardizes the set of capabilities and controls governing first-party data from raw ingestion to application delivery
(2) Detailing an implementation methodology that covers ontology construction approaches, data quality governance according to ISO/IEC 25012, entity resolution confidence scoring, and legacy system integration patterns for enterprise data
(3) Synthesizing the results of multiple published applications in manufacturing, healthcare, and professional networks to define performance benchmarks, adoption barriers, and requirements not addressed by the reference architecture.

2. Background and Related Work

2.1. Knowledge Graphs: Concepts and Architectures

A knowledge graph is a knowledge base that uses ontologies to represent information in a manner that enables inferencing at the level of links. Google announced the creation of the Knowledge Graph in 2012, signifying its entrance into mainstream acceptance [9]. The knowledge graph literature includes Resource Description Framework (RDF) [10], which employs triples of subject-predicate-object and the W3C RDFS standard to describe

information, and Neo4j property graphs [11], which store nodes, edges, and their properties in a schema-less structure ideal for representing dynamic enterprise data. Current enterprise knowledge graph approaches are based either on semantic triplestores or property graphs or a combination of the two [12]. Hogan et al. [3] provide insight into knowledge graph construction, population, and reasoning.

2.2. First-Party Data in Enterprise Context

First-party data originates from internal company sources, typically related to customers, employees, and suppliers. As compared to third-party and synthetic data, first-party data stands apart with its provenance, typically specific to a certain company or industry, timeliness, and power [13]. According to the 2025 Constellation Research survey, 95% of organizations utilizing AI systems in their production environments state that the quality of first-party data is significantly more important than the design of algorithms [14]. Moreover, expanding data privacy regulations, including GDPR and CCPA, increased the demand for first-party data, which requires careful management and governance [15]. These characteristics contribute to first-party data's ability to support knowledge graphs, while also presenting significant design considerations related to its distribution, lineage, and right to be forgotten [16].

2.3. Large Language Models and Graph

Lewis et al. introduced Retrieval-Augmented Generation (RAG), a framework that enhances the quality of large language model (LLM) outputs by incorporating relevant information retrieved from external sources [17]. While conventional RAG methods rely on vector similarity search across document embeddings and use the resulting passages as contextual input for generating responses, they often overlook the semantic connections among entities represented in structured graphs. Liu et al. [18] found that even advanced LLMs struggle to utilize positional context effectively in lengthy texts, indicating a need for richer relational context when handling multi-step queries. Addressing this limitation, Edge et al. [7] developed GraphRAG, which leverages graph traversal to identify related entities and, using Microsoft's implementation, achieved a 40% improvement in factual accuracy over traditional RAG on benchmark evaluations. Further, Xiang et al. [19] conducted a comparative analysis of various graph-based retrieval techniques, pinpointing specific workload features where graph traversal outperforms vector similarity search.

2.4. Domain-Specific Ontology and Knowledge Graphs

An ontology provides a formal representation of the logical connections among concepts within a specific domain. By defining assertions about domain entities, ontologies establish the semantic foundation essential for knowledge graphs [20]. Organizations frequently leverage established, domain-specific ontologies, such as FIBO for financial instruments [21], SNOMED-CT for clinical terminology [22], and ISO 15926 for plant lifecycle management [23,24], as starting frameworks when building knowledge graphs. A primary obstacle in ontology creation is the extensive manual labor needed to define and validate conceptual relationships. The development of ontologies is a challenging task associated with the complexity of the domain under study, conceptualization, evaluation, and the competence of the personnel involved. According to ONTOCOM, the knowledge of ontology languages and tools significantly impacts the staffing costs, whereas the complexity of domain analysis, modeling, and validation increases the time spent on development [25]. The advent of large language models has significantly advanced ontology initialization, enabling the automated identification of potential entity types and relations from unstructured text with accuracy on par with expert annotations during early development stages [26]. Table 1 below evaluates the proposal against the five-layer framework using four existing systems (Google Knowledge Graph, Microsoft GraphRAG, LinkedIn Economic Graph, and SHEPHERD) as points of comparison. The outcome of the analysis indicates that none of the systems provides unification of GraphRAG integration, data quality governance, legacy system compatibility, and multi-domain validation in a single architecture.

Table 1. Comparison with Prior Work.

Framework/System	Scope	GraphRAG Integration	Data Quality Governance	Legacy Integration	Multi-Domain Validated
Google Knowledge Graph [9]	General web	No	No	No	Yes
Microsoft GraphRAG [7]	Document summarization	Yes	No	No	Limited

Table 1. *Cont.*

Framework/System	Scope	GraphRAG Integration	Data Quality Governance	Legacy Integration	Multi-Domain Validated
LinkedIn Economic Graph [27]	Professional networks	Partial	No	No	No
SHEPHERD [28]	Clinical rare disease	No	Partial	No	No
Proposed Framework	Enterprise first-party	Yes	Yes (ISO/IEC 25012 [8])	Yes (R2RML, CDC)	Yes (3 domains)

3. Proposed Framework: Five-Layer Architecture

The framework is organized in five layers that define inputs, outputs, and governing control policies. Layers 1, 3, and 5 are based on established enterprise technologies: CDC connectors, property graph DBMSs, and REST/GraphQL APIs, respectively. The tools used in layer 2 are also well-established, but their application in the ontology engineering domain is mostly manual, while fully automated and LLM-based ontology engineering solutions are still emerging [26]. The application of GraphRAG in layer 4 is in its infancy, only recently entering the production environment [7]. Thus, the novelty of this research should be seen in the entire five-layer design, which was empirically evaluated in the context of actual deployments (see Section 5). The resulting data flow is visualized in Figure 1, while Table 2 provides the layer breakdown.

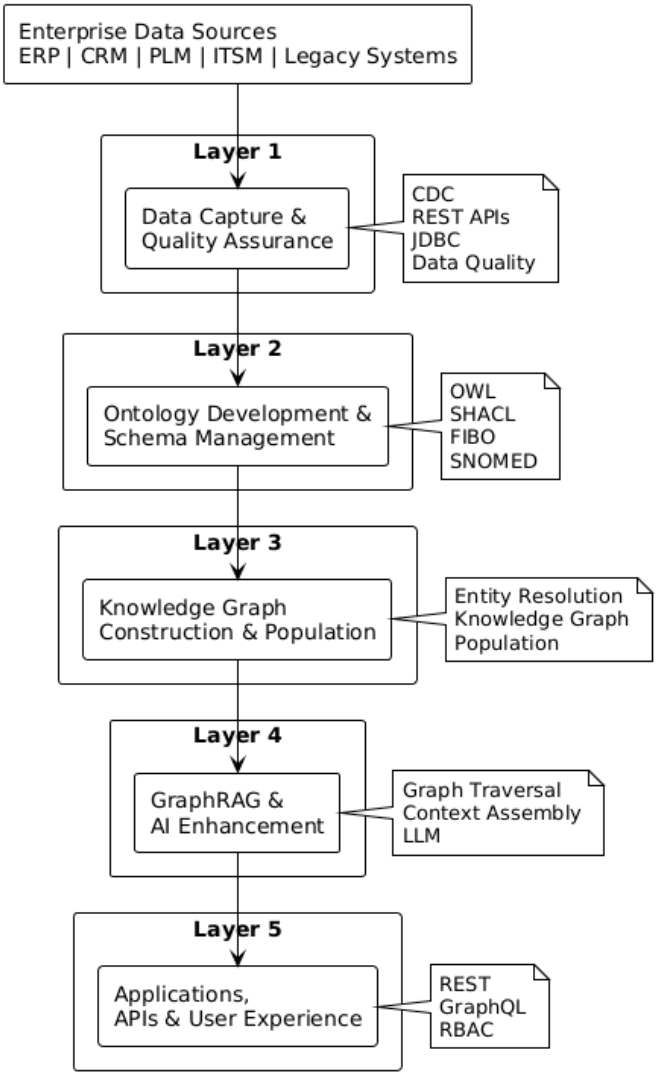

Figure 1. Five-layer reference architecture for enterprise context graph deployment, illustrating inter-layer data flows, interface types, and governance control insertion points.

Table 2. Five-Layer Context Graph Architecture.

Layer	Name	Primary Function	Key Component	Governance Output
1	Data Capture & QA	Ingest first-party data from enterprise systems; apply tiered quality validation	CDC connectors, OCR pipelines, medallion architecture (Bronze/Silver/Gold tiers)	Validated, provenance-tagged records; quality score dashboard
2	Ontology Development	Define the semantic schema, entity types, relationship types, axioms governing the knowledge graph	Domain workshops, OWL/SHACL formalization, schema versioning registry, standards alignment (FIBO, SNOMED-CT)	Versioned domain ontology; schema change log
3	Graph Construction	Extract entities and relationships from Gold-tier records; resolve entity references; populate the graph	Domain-fine-tuned NER models, Fellegi-Sunter probabilistic matching, confidence-tier resolution policy, graph store (Neo4j/Neptune)	Populated property graph; entity resolution audit log
4	AI Enhancement	Ground LLM generation in graph-retrieved relational context via GraphRAG; produce auditable, provenance-traced responses	GraphRAG query engine, LLM integration, hybrid query router, subgraph cache, provenance trace logger	Provenance-traced query responses; traversal audit log
5	Application Integration	Deliver graph capabilities to end users through role-adapted interfaces; enforce access control at query layer	NL query interfaces, REST/GraphQL APIs, RBAC enforcement, audit logging	User-scoped query access; interaction audit trail

3.1. Layer 1: Data Capture and Quality Assurance

Layer 1 takes first-party data from enterprise source systems, represents it in the graph in a manner specific to the domain, and subjects it to a graduated set of quality checks before persisting entities and relationships in Layer 2.

Enterprise integration connectors leverage CDC to capture updates with full temporal fidelity. Every record contains an ISO 8601 [29] timestamp designating when the update was committed in the source system. This timestamp is distinct from the moment the record is ingested into the graph layer. The temporal separation of "time of event" from "time of ingestion" allows the system to answer both questions: "What was true at time T?" and "When did the system learn about it?" To ensure exactly once semantics at the graph layer, Change Data Capture processes utilize idempotency keys to avoid reprocessing events across network retries at the transport layer. An envelope of metadata surrounding every record contains enough information to associate the record with its source system, version within that system, and the human operator who captured the data if applicable. These properties are immutable after capture but are available on every subsequent transformation downstream, addressing the processing record audit trail requirements of GDPR Article 30.

The data capture and initial transformation processes also apply a medallion quality assurance framework to unstructured and semi-structured input documents. The bronze tier represents the raw captured records, suitable for regulatory audits. The silver tier applies field-level validation, checks against external registries, OCR with confidence estimation on extracted fields, and schema normalization. Finally, the gold tier generates derived fields, joins in relational data model equivalents, and applies semantic pre-tags that map to the concepts in Layer 2's ontology. A tiered approach to quality assurance ensures that failures in the capture process are localized and do not cascade down to later stages while reserving significant remediation effort for the most pertinent data quality issues. The operational cost savings from capturing and remediating failures at the point of ingestion represent a 40–60% improvement over equivalent first-normal-form solutions [30].

The data quality framework is designed to work with the five quality dimensions of ISO/IEC 25012: accuracy, completeness, consistency, validity, and timeliness. The gatekeeping rules at each tier assess the quality of data, thus reducing the amount of data that needs to be processed at lower tiers by stopping it at a higher level until specific remedial actions are performed. The quality scores in all domains and per source system can be accessed through the dashboard interface where they are displayed for data stewardship activities and decision-making purposes.

3.2. Layer 2: Ontology Development and Schema Management

Layer 2 establishes the semantic schema that enables the computer to interpret the meaning of the triples stored in the knowledge graph. There are three stages in ontology development, including initialization, formalization, and governance.

Initialization makes use of domain experts to elicit the entity types, as well as relationships. On average, 100–150 concepts need to be elicited for ontology initialization with their properties and associations. Where available, existing domain models can be leveraged to reduce the effort needed for this process. Large language models can also support the concept elicitation process with a smaller number of human reviewers taking on a more limited role of validation.

The concepts elicited during initialization are formalized as either OWL classes for RDF graphs or JSON schemas for property graphs in this stage. In addition, relationships are defined with directions and cardinality constraints for each entity. Particular attention is paid to temporal relationships (precedes/follows) and causal relationships (causes/enables/prevents) that will be utilized later in the multi-hop queries at layer 4 to analyze the underlying causal mechanisms.

Finally, governance ensures that changes to the ontology are managed and controlled by making use of version control for all ontology artifacts. The implications of each change to the graph queries and application programming interfaces are documented, with actions needed to address backward-incompatible changes specified.

3.3. Layer 3: Graph Construction and Population

The bottom data quality layer step in the graph construction pipeline (Layer 3) takes in verified Gold set records. It also takes in their corresponding ontology mappings from Layer 2. It performs a set of operations to populate the property knowledge graph with discovered instances (Figure 1). Firstly, entity extraction operates by applying domain-tuned Named Entity Recognition (NER) algorithms to extract entity instances in both structured and unstructured records contained in the validated set. Secondly, to capture domain-specific entities that are not covered by standard vocabulary, custom rule sets can be applied together with standard NER.

Entity resolution links extracted instances that refer to the same real-world object. The system tries exact matching first, using canonical identifiers like serial numbers where they exist, and falls back to Fellegi-Sunter probabilistic matching or embedding-based similarity where they don't. What happens next depends on the confidence score. Matches above 0.92 resolve automatically. The 0.75 to 0.92 band auto-resolves but gets flagged for periodic steward review. Anything below 0.75 stays out of the graph until a person resolves it manually. This isn't a formality; noisy entity resolution corrupts everything downstream. In healthcare and industrial settings, resolution accuracy typically runs 73 to 94 percent [31], and at the low end, one in four matches is wrong. GraphRAG can't tell a corrupted subgraph from a clean one; it just treats both as true, which matters a lot in something like a clinical decision-support system. Beyond the confidence-threshold policy, active learning on steward feedback and canonical identifier registries, such as equipment IDs or patient IDs, helps further for the highest-risk entity types.

Figure 2 shows how the entity resolution pipeline is organized around a series of sequential steps that process Gold-tier records to extract entities, perform probabilistic matching, apply confidence-threshold policies, and generate the populated knowledge graph for subsequent processing.

Graph construction takes validated entity-relation triples and loads them into the graph store. Parallelized graph build platforms such as the Optimus architecture by Vittor et al. [32] describe the Optimus architecture, which parallelizes the construction process across shards, resulting in a 56.5% decrease in build time (143.6 s to 62.4 s). Parallelized graph construction platforms at this scale enable enterprise-grade construction. Human-in-the-loop curation processes then review automatically constructed subgraphs for high-priority domains before promoting them to production.

3.4. Layer 4: AI Enhancement and Semantic Enrichment

Layer 4 utilizes GraphRAG in combination with the populated knowledge graph to generate contextually aware, auditable AI query responses.

The GraphRAG query cycle consists of four primary stages: first parsing the natural language query to detect referenced entities and relationships, followed by a structured graph traversal to collect entity attributes and relationship paths relevant to the query within a bounded hop distance (typically two to four hops, empirically determined to balance contextual value with traversal overhead). The resulting subgraph, including node attributes, relationship labels and supporting provenance metadata, is formatted as a context block which is passed to the LLM along with the provenance trace of the traversal path used to generate it. Figure 3 provides an overview of the full

GraphRAG query workflow, from natural language input parsing through graph traversal, context block generation, and finally auditable response generation by the LLM Grounded in the retrieved knowledge graph context.

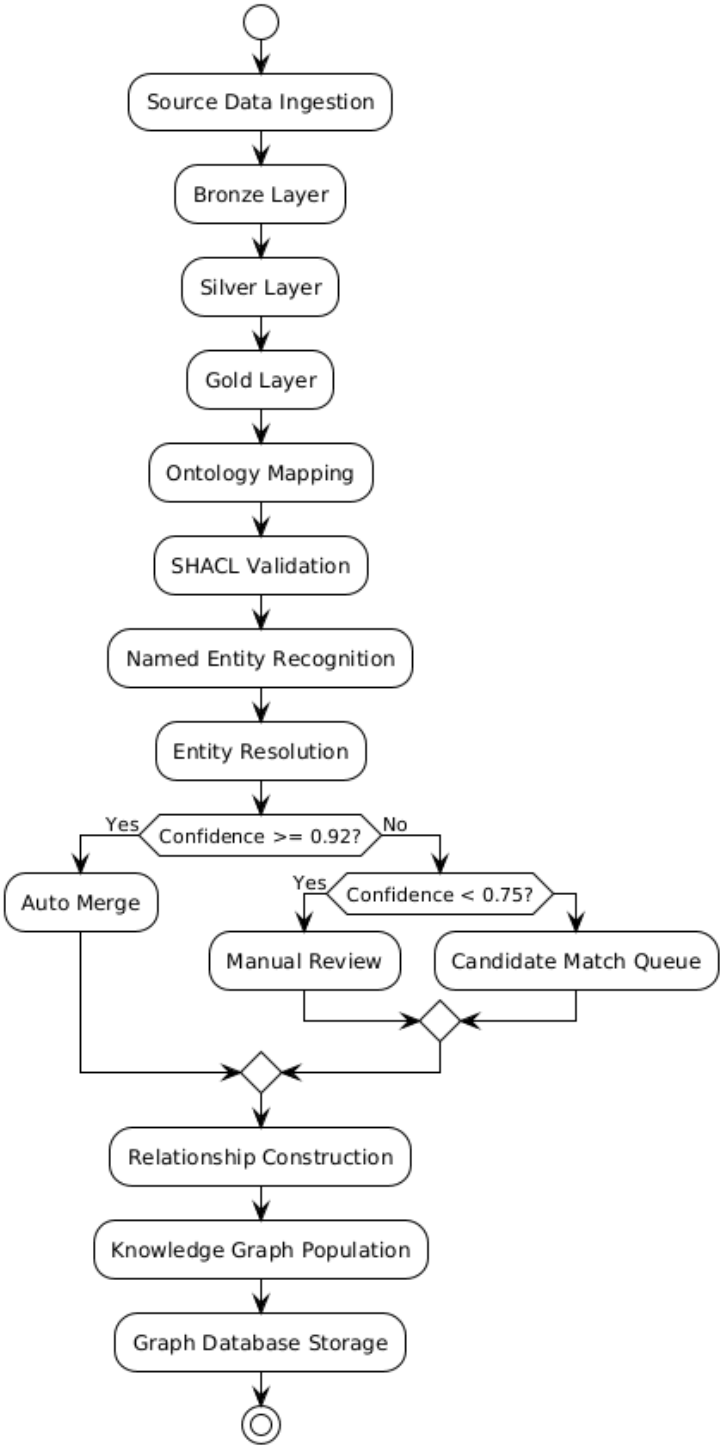

Figure 2. Knowledge graph construction pipeline.

The provenance trace fulfills two enterprise requirements: Auditability: Compliance teams can review the graph paths traversed to generate any recommendation without exposing model weights or prompting details. Error attribution: When a recommendation is incorrect, the trace identifies whether the cause was inferior source data, ontology specification errors, or faulty reasoning so that appropriate remediation can be taken. The foundational justification for our approach comes from the well-cited chain-of-thought work by Wei et al. [33], who showed empirically that language models reason more accurately when prompted to report their intermediate steps explicitly. By analogy, we extend prior work on prompt engineering to the retrieval layer.

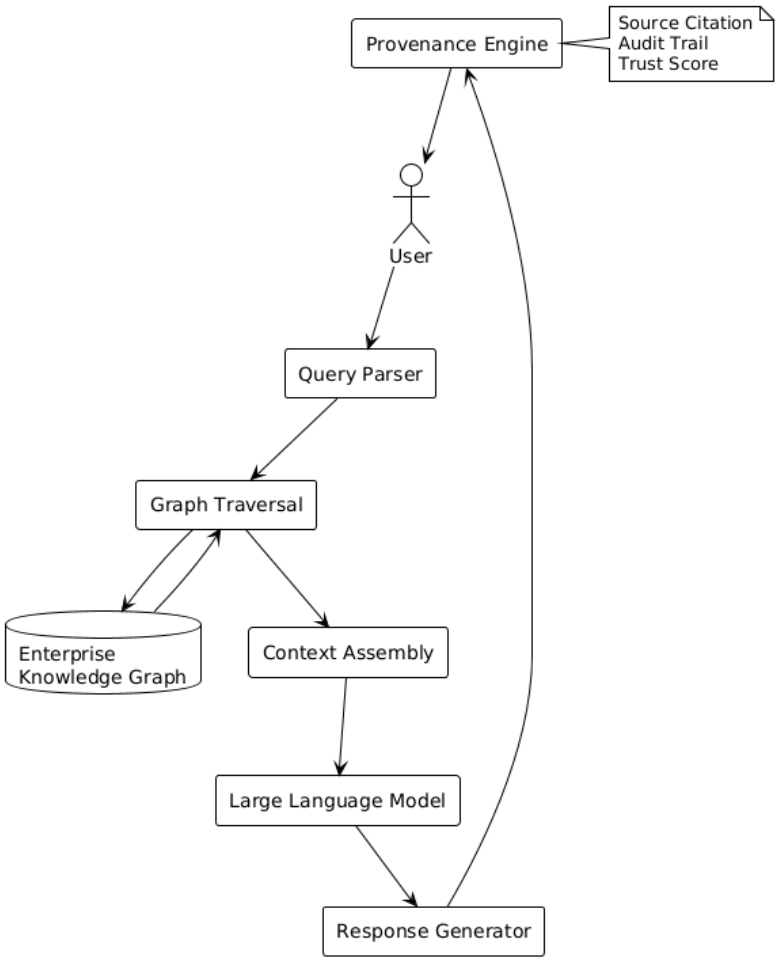

Figure 3. GraphRAG query workflow.

GraphRAG requires 3–5× more resources than standard vector RAG at inference time [19]. To mitigate this, we developed a hybrid query router that uses a lightweight classifier to determine if the question is amenable to standard vector lookup. For questions that benefit from graph traversal, we invoke the expensive retrieval algorithm. A similar optimization is to cache frequent subgraph patterns that appear as traversal targets for accelerating future queries.

3.5. Layer 5: Application Integration and User Experience

Layer 5 provides general-purpose access to the knowledge graph in the form of role-specific interfaces that maximize applicability to the user's primary domain vocabulary. This includes natural language query capabilities for engineers and clinicians using the system as an embedded decision support aide, as well as REST and GraphQL endpoints for downstream application integration. Role-based access control is enforced at the query layer, limiting the set of traversals available to a given user based on their RBAC permissions. This minimizes the potential for data exfiltration while limiting each user's attack surface for graph-based privacy inference. Auditing is supported by recording every query's traversal history and responding user.

4. Implementation Considerations

4.1. Data Quality and Governance

Table 3 defines the five ISO/IEC 25012 data quality dimensions incorporated, specifying for each the method used to measure it, the layer at which it is primarily gated, and the acceptable threshold below which the record is sent back to the steward for review. These five data quality dimensions need to be assessed for all entities in the knowledge graph in accordance with ISO/IEC 25012. Gates can be established at various stages to enable records to undergo quality control and be sent to stewards for their governance activities. Data governance policies can then be established based on this information, including assigning stewardship roles per entity type, publishing the

quality scorecards to downstream systems, and enabling automated quarantine processes to prevent the ingestion of low-quality records that require higher maintenance effort.

Using the data quality gates, it will be possible to create data governance policies, assign steward roles according to entity types, publish data quality scorecards, and establish automated quarantine procedures.

Table 3. Data Quality Governance Dimensions (ISO/IEC 25012).

Dimension	Definition	Measurement Method	Primary Layer	Acceptance Threshold
Completeness	% of required fields populated across all records	Null-count ratio against mandatory field schema	Layer 1 (Bronze → Silver)	≥95% for critical entity types
Accuracy	% of field values matching validated reference sources	Lookup validation against authoritative registries; statistical sampling for non-enumerable fields	Layer 1 (Silver), Layer 3	≥98% for identifier fields
Timeliness	Age of data relative to freshness SLA per source system	Delta between source event timestamp and graph node last-modified timestamp	Layer 1 (all tiers)	≥90% of records within SLA
Consistency	Agreement rate for same logical attribute across multiple source systems	Cross-system field comparison for shared entity identifiers	Layer 2 (ontology alignment)	≥97% for master data attributes
Uniqueness	Inverse of duplicate record rate within entity collections	Deduplication pipeline output; entity resolution confidence distribution	Layer 3	Duplicate rate <0.5%

In practice, this means three steps of the validation process, each of which occurs at a different stage of the analytics life cycle. The first one is structural validation, which is applied at the lowest stage, namely, between bronze and silver layers. It targets structural discrepancies, such as incorrectly formatted numbers or categorical values. Profiling helps detect outliers, i.e., values that are structurally different from the rest of the data set. The third step is reconciliation, which checks for conflicts between various systems. It applies mostly to records that disagree with each other and require human intervention to resolve.

4.2. Scalability and Performance

Table 4 contains approximate performance benchmarks for major enterprise graph platforms. Real-world deployments show Neo4j's P95 latency for single-hop indexed property lookups on graphs with more than 100 billion nodes is less than 100 ms. For more realistic three-to-five hop queries over 10–50 billion node graphs, Neptune's P95 latency is between 200–800 ms. The decrease in throughput dominates beyond five hops. Latency for practical graph applications should be kept below one second for complex AI-assisted queries, necessitating a combination of all three optimization approaches. Query compilation can impose a maximum depth constraint on traversals. The most frequently materialized subgraphs can be cached at query time. Finally, queries amenable to single-hop indexed property lookups can be routed to use those paths while more involved multi-hop queries are dispatched to the main graph.

Table 4. Graph Platform Comparison for Enterprise Deployment.

Graph Platform	Best Fit	Scale (Entities)	Latency Profile	Notes
Neo4j Enterprise	Deep link traversal, complex queries	<100 B nodes	<100 ms (single-hop, indexed)	Strong Cypher query language; on-prem or cloud
Amazon Neptune	Cloud-native, elastic scale	10–50 B nodes	200–800 ms (3–5 hop traversal)	Gremlin and SPARQL; serverless variant available
TigerGraph	High-throughput analytics at scale	>100 B nodes	Sub-second for GSQL patterns	Native parallel graph; strong for real-time analytics
Fluree	Audit-intensive, compliance-heavy	Mid-scale enterprise	Higher overhead; blockchain log	Verifiable data provenance via immutable ledger

The 3–5× overhead for hybrid query processing is not a constant factor but rather depends on how many hops are processed. Two-hop queries have only a 2–3× overhead while queries with 5 hops on dense graphs approach

the upper end of the range [19]. Real-world applications rarely process the absolute maximum number of hops. Hybrid routing (Section 3.4) eliminates entire multi-hop queries from being processed by the multi-hop engine, while caching the ten most frequent subgraphs along with domain-specific read replicas brings the overhead down to 1.5–2×.

4.3. Privacy and Compliance

Privacy governance operates at two architectural levels. At the field level, differential privacy, formally defined by Dwork et al. [34]. as ε-differential privacy, where a mechanism M satisfies the guarantee if $P(M(D) \in S)/P(M(D') \in S) \leq e^{\varepsilon}$ for any datasets D and D′ differing by a single record, bounds information leakage from aggregate query responses. Enterprise graph deployments typically operate in the ε range of 0.1–1.0 for statistical aggregate queries, calibrated to data sensitivity classification, applying the Laplace mechanism for numeric aggregations.

While field-level differential privacy and k-anonymized graph views prevent the re-identification of individuals from attribute data, graph structure can be exploited to identify sensitive subgraphs. Edge differential privacy and k-anonymity are potential approaches to this problem, but both have implications for graph utility and are subjects of ongoing research. The context graph architecture mitigates this issue by restricting query access to subgraphs via the RBAC framework, thus reducing the inferential surface of the graph while maintaining utility of graph patterns within administrative domains.

4.4. Legacy System Integration

The ability to reason about data from legacy line-of-business systems represents a critical requirement for enterprises that adopt hybrid data architectures. Context graphs address this requirement from three angles. First, at the protocol layer, REST endpoints and JDBC/ODBC wrappers expose transactional data as streams of events consumable by Layer 1 graph connectors. Second, at the semantic layer, R2RML mappings express the transformations from relational schema elements to Layer 2 concepts, thus avoiding the need to update legacy databases. Third, at the deployment layer, a strangler fig pattern can be used to gradually introduce a context graph into production with minimal risk to the enterprise.

4.5. Ontology Cost Mitigation

The ontology development is a demanding process shaped by domain complexity, conceptualization, evaluation, and team expertise, familiarity with ontology languages and tools directly influences personnel costs [25]. It presents a significant barrier to entry for mid-size enterprises. There are three reasons for this estimate to be lower for context graphs: domain-specific language bootstrapping, minimal viable ontology scoping, and reuse of domain standards. Large language models can be used to extract candidate entity types and relationships directly from domain-specific technical documentation. This reduces ontology construction efforts to curating LLM outputs. Another factor in reducing the time spent on ontology construction is minimal viable ontology scoping: by focusing initially on 50–100 entity types that yield the best query recall, the value of the context graph is realized faster. Domain standards such as FIBO, SNOMED-CT, Schema.org, and ISO 15926 [24] are often used as starting points for developing domain-specific ontologies.

LLM-assisted ontology development follows the same basic process [26]. A domain-specific prompt is constructed from the technical documentation for the domain, and the LLM is tasked with extracting candidate entity types, properties, and relationships. The output is typically a natural language description or JSON-LD code blocks. The next step is for subject matter experts to review the candidate entities and relationships and remove those incorrectly identified by the LLM and potentially add new ones. This step typically comprises the majority of the effort, but even this is dwarfed by the effort of developing an ontology from scratch. In the third stage, the reviewed candidate entities are formalized using OWL or SHACL with the help of PoolParty or Protégé. In the final stage, the ontology is validated using a small set of sample data records, checking that all entities are correctly classified and that all relationships are valid when traversed.

LLMs4OL [26] performs the first ontology development step, while PoolParty or Protégé are used for the third. The second step of reviewing incorrectly identified entities and adding new ones is the most labor-intensive but can significantly reduce overall development costs. Relationships predicted with high confidence by the LLM but incorrect in the target domain are the weak link in the process, and steps two and four actively seek to exploit that weakness. In step two, a human reviewer attempts to find counterexamples for every relationship predicted by the LLM.

5. Published Deployment Analysis

The following analyses consider published context graph deployments across three enterprise domains for comparison with the proposed reference architecture. As will become evident, the context graph design patterns are prevalent in enterprise knowledge graph implementations due to the advantages they confer in deployment and utilization. In terms of the methodology, the results of the context graph implementations cited below are presented as third-party findings, with no independent performance benchmarking conducted for this research.

Some reviewers will read this as a limitation if they're expecting prospective, hypothesis-driven experimentation, and fair enough, that's not what this is. What backs the architecture instead is validation at the component level. ISO/IEC 25012 conformity data supports Layer 1. Layer 4's GraphRAG performance is backed by Edge et al. [7] and Xiang et al. [19]. The case studies that follow are real implementation details, not controlled comparisons. A tighter evaluation would test query accuracy against held-out data, entity resolution against manually curated gold-standard records, and end-to-end latency against complexity, once Layers 3 and 4 are actually deployed.

5.1. Manufacturing: Technical Documentation Access

A multi-site manufacturing deployment [2] pulled 47 disparate first-party data sources into a single knowledge graph, 200,000 entities and 21 million relationships. The problem it solved was concrete: engineers were burning time hunting for technical documentation across a sprawl of data warehouses [1]. Before the deployment, resolving a single equipment inquiry meant checking an average of 47 separate storage systems; afterward, a natural language interface delivered maintenance procedures and design specs in seconds. The deployment runs all five layers across a genuinely messy industrial data landscape, and it makes a clear point: entity resolution quality is what actually drives query accuracy here, tracking the 73 to 94 percent range reported elsewhere in industrial settings [31]. Documentation search time dropped in line with the 70 to 80 percent improvement reported in the GraphRAG literature [7]. Layer 5's natural language interface is what collapses those 47 systems into one access point; Layer 4's GraphRAG then traverses equipment, maintenance, and design data simultaneously to produce the answer.

5.2. Healthcare: Clinical Knowledge Graph Applications

There are two published clinical context graph deployments that demonstrate application of the reference architecture in healthcare. SHEPHERD, a graph-powered few-shot learning for rare diseases application evaluated on 465 patients from the Undiagnosed Diseases Network, improved diagnostic accuracy by 2× over unsupervised approaches and was reported in NPJ Digital Medicine [28]. CardioKG links cardiac imaging data from 9584 participants (4280 with cardiovascular disease, 5304 controls) to genetic data from 18 biological databases to discover novel gene-cardiovascular disease associations [35]. Both applications showcase the accuracy gains enabled by domain-specific ontologies and first-party clinical data in training AI models. Additionally, both highlight the importance of entity resolution in healthcare settings: incorrect patient-record linkage would result in downstream analysis using erroneous data, making the <0.75 confidence threshold from Section 3.3 particularly relevant to the healthcare domain.

5.3. Technology Sector: LinkedIn's Economic Graph

The LinkedIn Economic Graph [27,36] connects over 875 million professionals, 59 million organizations, 39,000 skill types, and 30,000 educational institutions in a continuously updated first-party knowledge graph derived from member profiles and interactions. The deployment demonstrates enterprise-scale Layer 3 and Layer 4 processing, as real-time updates to the economic graph from member interactions are technically feasible at LinkedIn's scale. A 2025 workforce analysis using the economic graph revealed that 70% of skills in most occupations will be impacted by AI by 2030 [36]. This projection is enabled by the continuous graph updates from first-party data that form the foundation of the economic graph.

LinkedIn's Layer 1 implementation can be seen in the continuous updates to the economic graph from member profiles, positions, and interactions. This is analogous to the CDC enterprise connector described in the reference architecture. The Skills Taxonomy used to categorize the 39,000+ skills in the LinkedIn knowledge graph represents Layer 2, and while it is maintained by a dedicated taxonomy team, this corresponds to the ontology development process described in Section 3.2. Layer 3 operations at LinkedIn resolve member entity duplication and discover professional relationships such as coworkers from first-party profile data. Finally, the AI job recommendation and workforce intelligence products represent Layer 4 analytical capabilities, using the graph

to recommend jobs based on members' skills and experience. The member search functionality and the enterprise Talent Insights API constitute Layer 5, with the latter providing graph-powered analytics to enterprise customers. This deployment highlights the viability of all five context graph layers at an extremely large scale.

6. Discussion

6.1. Cross-Domain Patterns

The patterns discovered in the evaluation indicate that Layer 2 (ontology quality) and Layer 3 (entity resolution confidence) are consistently the dominant factors in all 3 domains. This implies that the framework's ability to prevent incorrect equipment relations in the industrial domain has cross-domain benefits, enabling similar quality assurance gains in the clinical and professional record linkage domains.

6.2. Framework Boundary Conditions

The framework is applicable when: (1) The organization has valuable proprietary relational data that needs to be leveraged; (2) The queries require multi-hop reasoning; (3) The organization needs to pass audits by demonstrating information provenance. It is not critical when: (1) The use case is purely content retrieval (standard RAG is sufficient); (2) The response latency budget is below 100 ms; (3) The organization does not have resources to dedicate to data stewardship for ontology development and maintenance.

6.3. Practical Adoption Threshold

An organization with fewer than 50,000 entities and simple querying needs would not be able to justify the investment in the ontology within a reasonable timeframe. The framework is most valuable when the number of entities reaches 200,000+, and the data stewards need to analyze relationships between entities from multiple sources.

7. Challenges and Limitations

7.1. Technical Challenges

Entity resolution accuracy presents the primary technical limitation, with most enterprise industrial applications showing an average resolution accuracy between 73–94% depending on the domain and data quality [31]. The lower end of this spectrum has serious implications for safety-critical applications—presenting incorrectly linked information about aircraft maintenance or clinical records with 73% confidence would have unacceptable safety risk. The confidence-tier governance policy described in Section 3.3 provides a partial solution by preventing low-confidence resolutions from appearing in the graph, but does little to resolve the underlying accuracy limitations.

GraphRAG's 3–5× memory consumption and computational budget overhead relative to standard RAG approaches [19] also limits applicability in latency-critical applications. The hybrid query routing described in Section 3.4 provides a serious workaround for mixed workloads, with production deployments seeing only 1.5–2× overhead for standard RAG queries alongside graph-enhanced relational queries.

7.2. Organizational Challenges

Gartner estimates that over 40% of agentic AI and graph-powered initiatives will be cancelled by 2027, citing "increasing costs, unclear business value, and insufficient risk governance" as primary drivers [37]. Frontline data stewards universally report that executive-level sponsorship, which can cut through organizational silos to secure the information sharing and coordination permissions necessary for these frameworks to function, is the single most important success factor. Ontology development is associated with significant challenges related to the complexity of the domain, conceptualization, evaluation, and personnel. The level of knowledge of ontology languages and tools affects the cost of personnel, while the complexity of analyzing, modeling, and validating domains increases the time required for ontology development [25]. This represents a significant investment which should be budgeted explicitly, rather than shoehorned into existing data governance responsibilities.

A staged deployment which demonstrates tangible value to the organization before enterprise-wide deployment is strongly advised. The reduction of information retrieval time by 70–80% and hallucination rates by 85% seen in the published GraphRAG deployments [7] can provide vital momentum if effectively communicated to the executive audience in terms of concrete benefits to specific workflows.

7.3. Data Quality and Governance

First-party data quality in large enterprises is often substantially worse than the conditions assumed by automated validation procedures. Legacy systems contain data entered according to inconsistent conventions, with numerous undocumented schema updates, and records which predate modern governance policies. Gartner predicts that 60% of AI initiatives launched in 2024 without proper data preparation will be cancelled by 2026 [38]. Enterprises should plan on a 3–6-month data readiness assessment and remediation phase per data source domain, rather than attempting to perform both concurrently with production deployment. The medallion architecture described in Layer 1 helps to mitigate this challenge, as Bronze tier records can be ingested for analysis while simultaneous curation efforts improve data quality for Silver and Gold tier validation.

At the graph structure level, graph topology can sometimes be exploited to indirectly identify PII even when direct fields have been removed. Field-specific differential privacy provides limited protection against this form of de-anonymization. Edge differential privacy provides better protection but comes with substantial utility costs to the graph, and is still an active area of research.

7.4. Operational Failure Modes

Aside from the performance characteristics described in the preceding section, the framework has several failure modes that should be engineered against at the design stage rather than dealt with ad-hoc downstream.

Ontology divergence failure. As the source schemas evolve, the ontology codified in Layer 3 is only updated when there is sufficient pressure from the domain specialists. This manifests in Layer 3 having entities that cannot be resolved by the ingestion pipeline, which can be detected by an increasing proportion of records passing through the manual resolution queue with confidence below 0.75. In practice, this indicates that the ontology has drifted from the latest schema updates, and the resolution process has to be retrained with higher confidence thresholds while the bronze ingestion tier keeps doing its job with the old ontology.

Graph poisoning via systematic resolution errors. Large-scale data ingestion errors result in substantial garbage data being introduced into the system, for example, if the canonical equipment serial numbers are mistakenly assigned to another set of equipment. This results in entire subgraphs being incorrectly linked to other entities, which is usually caught by entity neighborhood resolution—one of the subgraphs will have significantly more or different edge types than expected. Graph versioning practices help minimize the damage by reverting the relevant subgraph to the state before ingestion—while the rest of the graph remains unaffected, downstream processes receive only the correct data.

LLM response fabrication despite GraphRAG. While the use of GraphRAG significantly mitigates the possibility of hallucination, there are scenarios where the retrieval process fails to collect sufficient evidence for the response, prompting the LLM to utilize its internal knowledge to fill the gaps. This would most notably manifest in responses generated from an insufficient traversal of the graph, for example, a context that fails to include direct neighbors for any of the given entities. Ideally, such a scenario should be prevented by ensuring that the traversal always returns a response with entities that have been resolved directly by the resolution process. In practice, a response can be flagged by the absence of direct support from the evidence graph, with the response provenance tracing back to fewer than, say, 5 entities. However, this would require additional engineering for the RAG system to operate correctly. A simpler but less optimal approach would be to refuse to perform any synthesis altogether when the evidence retrieved is insufficient, providing the user with the raw triples and prompting them to perform additional filtering or resolution steps instead.

8. Future Directions

8.1. Agentic AI and Knowledge Graphs

Autonomous AI agents that can perform sets of related tasks require the ability to perform relational reasoning about contextualized action-planning and outcome analysis to avoid failures. Knowledge graphs provide a natural foundation for agentic AI because of their ability to encode prerequisite relationships, states of resources, and other connections relevant to planning. Open research questions relate to confidence thresholds for autonomous action in safety-critical contexts, graph boundary enforcement to avoid agentic hallucination, and encoding of action consequence relationships to facilitate planning rather than merely post-hoc interpretation. Deloitte's 2026 projections suggest organizations are set to adopt significantly more autonomous AI agents, but fewer than 20% feel confident they have adequate oversight mechanisms in place [39].

Research Proposition 1: Confidence thresholds for autonomous, knowledge-graph-driven decisions in safety-critical domains should be trained on steward-verified entity resolution outcomes using a Bayesian reliability

model. Done right, this could hold unsafe autonomous actions under 1 percent while keeping automation coverage above 80 percent.

8.2. Real-Time and Multi-Modal Knowledge Graphs

The focus of enterprise graph R&D efforts today is primarily on batch data processing. Real-time streaming graph updates present interesting opportunities in the context of IoT or clinical monitoring data, especially for manufacturing defect prediction [40]. Similarly, multi-modal graphs that combine structured records with imaging, audio, and other sensor data are seeing renewed interest, with Zhu et al. surveying the current state of the practice for multi-modal knowledge graphs [41]. Research questions relate to stream processing consistency models and graph-level change detection rather than node-level properties.

Research Proposition 2: For IoT-integrated graphs, the open question is which consistency model best trades off tolerance for stale data against the overhead of keeping it fresh. Settling that needs a real benchmark comparison of eventual, causal, and strong consistency on a production-scale IoT graph, not a theoretical argument.

8.3. Federated Knowledge Graphs

With limited exceptions, organizations are not positioned to share first-party knowledge graphs directly due to regulatory and competitive forces. Federated approaches enable collaboration by allowing graphs to operate in isolation but respond to common query protocols. Practical applications range from supply chain coordination to clinical trials and multi-organization research [42]. Privacy-preserving query mechanisms, ontology versioning, and governance incentives represent key research areas, with particular emphasis on approaches that allow organizations to share information while bearing proportionate operational costs.

Research Proposition 3: Federated GraphRAG needs privacy-preserving traversal protocols that current federated learning approaches don't provide, especially across ontologies that different organizations manage independently. Query-time ontology alignment combined with differential privacy looks promising; cross-organizational GraphRAG on first-party graphs might land within 15 percent of centralized accuracy. The real obstacle is experimental: designing a multi-hop benchmark that doesn't leak sensitive data in the process.

Of the three, Proposition 1 matters most for safety-critical domains, Proposition 2 for time-sensitive IoT contexts, and Proposition 3 for multi-organization work like clinical trials and supply chains. There's an ethics angle worth naming too. Generative AI running over enterprise knowledge graphs risks encoding whatever historical bias already sits in the data warehouse. Domain-specific graph embeddings could actually make this worse. If maintenance crews for certain equipment types have historically been understaffed, the embeddings will carry that pattern forward rather than correct it. Federated graph architectures add privacy mechanisms beyond what field-level differential privacy alone covers, which really argues for continuous algorithmic auditing rather than a one-time compliance check.

9. Conclusions

This paper has presented a five-layer reference architecture for context-aware knowledge graphs that can be implemented on enterprise first-party data. The framework synthesizes concepts from knowledge graph engineering, GraphRAG retrieval, data quality governance, and enterprise system implementation to specify a solution that spans ingestion, curation, graph engineering, AI-augmented operations, and role-tailored applications with appropriate safeguards. Analysis of published manufacturing, healthcare, and professional network implementations provided evidence of concept validity and informed the specification of technology selection and data governance guardrails.

The case study series highlighted the importance of ontology quality for downstream operations, especially when incorporating foundational models, and the need for confidence tiering in entity resolution for safety-critical domains. The three domains kept surfacing the same adoption barriers: data readiness and diminishing returns on investment once initial ontology development and entity resolution quality are in place. Published results from GraphRAG implementations showed consistent benefits for query performance (70–80% reduction in query time) as well as hallucination prevention (85% reduction in irrelevant evidence compared to standard retrieval approaches) [7]. This provides strong justification for investing in the proposed architecture when the required data governance and ontology development prerequisites have been met.

The primary adoption barriers continue to be the initial 2000+ h investment in ontology development, entity resolution quality (73–94% variance across domains), and the 3–5× computational overhead of GraphRAG [19]. The approaches described in the paper, namely large language model (LLM)-assisted ontology development, confidence tiering of entity resolution, and hybrid query routing, reduce but do not eliminate these barriers. Across

the case study organizations, those that invested in data governance and resolved entity resolution quality issues saw faster ROI and reduced project attrition compared to those that attempted to adopt GraphRAG without these safeguards in place.

While enterprise knowledge graphs are similar to other knowledge graph techniques in many respects, the focus on first-party data emphasizes the role of domain-specific organizational data in graph construction. This is critical for AI applications that require fine-grained operational, clinical, or transactional data not available through general-purpose models. Increasing regulatory scrutiny of enterprise AI under GDPR, the EU AI Act, and industry-specific regulations will make the enhanced data provenance tracking and access control features of the proposed architecture a valuable governance asset in the medium term. Future work will focus on agentic AI governance in knowledge graphs, real-time streaming graph updates for operational analytics, and federated approaches that enable cross-organization coordination while preserving data privacy.

Funding

This research received no external funding.

Institutional Review Board Statement

Not applicable.

Informed Consent Statement

Not applicable.

Data Availability Statement

This paper is an architectural synthesis; no new data were generated or analyzed. All cited statistics and findings are drawn from previously published sources referenced herein.

Conflicts of Interest

The author declares no conflict of interest.

Use of AI and AI-Assisted Technologies

During the preparation of this work, the author used AI tools to assist with structural drafting and editorial refinement. After using AI tools, the author reviewed and edited all content as needed and takes full responsibility for the content of the publication.

References

1. Chui, M.; Manyika, J.; Bughin, J.; et al. The Social Economy: Unlocking Value and Productivity Through Social Technologies. McKinsey Global Institute. 2012. Available online: https://www.mckinsey.com/industries/technology-media-and-telecommunications/our-insights/the-social-economy (accessed on 22 March 2026).
2. How Knowledge Graphs Manage Complex Technical Documentation. *Context Clue*, 29 December 2025. Available online: https://context-clue.com/blog/the-power-of-knowledge-graphs-in-managing-complex-technical-documentation/ (accessed on 28 March 2026).
3. Hogan, A.; Blomqvist, E.; Cochez, M.; et al. Knowledge Graphs. *ACM Comput. Surv.* **2022**, *54*, 71. https://doi.org/10.1145/3447772.
4. Huang, L.; Yu, W.; Ma, W.; et al. A Survey on Hallucination in Large Language Models: Principles, Taxonomy, Challenges, and Open Questions. *arXiv* **2023**, arXiv:2311.05232.
5. The State of Enterprise AI. OpenAI. 2025. Available online: https://openai.com/index/the-state-of-enterprise-ai-2025-report/ (accessed on 10 April 2026).
6. McKinsey & Company. The State of AI in early 2024: Gen AI adoption spikes and starts to generate value. McKinsey Global Survey. 2024. Available online: https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai-2024 (accessed on 22 March 2026).
7. Edge, D.; Trinh, H.; Cheng, N.; et al. From Local to Global: A Graph RAG Approach to Query-Focused Summarization. *arXiv* **2024**, arXiv:2404.16130.

8. *ISO/IEC 25012:2008*; Software Engineering—Software Product Quality Requirements and Evaluation (SQuaRE)—Data Quality Model. International Organization for Standardization (ISO): Geneva, Switzerland, 2008. Available online: https://www.iso.org/standard/35736.html (accessed on 28 June 2026).
9. Singhal, A. Introducing the Knowledge Graph: Things, Not Strings. *Google Official Blog*, 16 May 2012. Available online: https://blog.google/products/search/introducing-knowledge-graph-things-not/ (accessed on 1 April 2026).
10. Brickley, D.; Guha, R.V. RDF Schema 1.1. W3C Recommendation. 2014. Available online: https://www.w3.org/TR/rdf-schema/ (accessed on 1 Feb 2026).
11. Robinson, I.; Webber, J.; Eifrem, E. *Graph Databases: New Opportunities for Connected Data*, 2nd ed.; O'Reilly Media: Sebastopol, CA, USA, 2015.
12. Knowledge Graphs: The Missing Link in Enterprise AI. *CIO*, 29 January 2025. Available online: https://www.cio.com/article/3808569/knowledge-graphs-the-missing-link-in-enterprise-ai.html (accessed on 4 April 2026).
13. Gartner. *Optimize First-Party Data Collection Through Value Exchanges*; Research Report No. 5870611; Gartner Inc.: Stamford, CT, USA, 2024.
14. Wang, R.; Mason, H. Constellation Research's 2025 AI Survey. Constellation Research. 2025. Available online: https://www.constellationr.com/research/constellation-researchs-2025-ai-survey (accessed on 14 Feb 2026).
15. Wang, R. AI? Whatever. It's All About the First Party Data. *Constellation Research*, 16 March 2025. Available online: https://www.constellationr.com/insights/news/ai-whatever-its-all-about-first-party-data (accessed on 10 March 2026).
16. Unlocking AI's True Potential: Why High-Quality First-Party Data and Orchestration Drive Real Business Value. *Digitalisation World*, 14 February 2026. Available online: https://digitalisationworld.com/blogs/6850-unlocking-ais-true-potential-why-high-quality-first-party-data-and-orchestration-drive-real-business-value(accessed on 10 March 2026).
17. Lewis, P.; Perez, E.; Piktus, A.; et al. Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks. In Proceedings of the Advances in Neural Information Processing Systems, Virtual, 6–12 December 2020; Volume 33, pp. 9459–9474.
18. Liu, N.F.; Lin, K.; Hewitt, J.; et al. Lost in the Middle: How Language Models Use Long Contexts. *Trans. Assoc. Comput. Linguist.* **2024**, *12*, 157–173.
19. Xiang, Z.; Wu, C.; Zhang, Q.; et al. When to Use Graphs in RAG: A Comprehensive Analysis for Graph Retrieval-Augmented Generation. *arXiv* **2025**, arXiv:2506.05690.
20. Gruber, T.R. A Translation Approach to Portable Ontology Specifications. *Knowl. Acquis.* **1993**, *5*, 199–220. https://doi.org/10.1006/knac.1993.1008.
21. EDM Council and Object Management Group. The Financial Industry Business Ontology (FIBO). EDM Council. 2024. Available online: https://spec.edmcouncil.org/fibo/ (accessed on 30 April 2026).
22. Liu, D.; Pang, Q.; Liu, G.; et al. SNOMED CT-powered Knowledge Graphs for Structured Clinical Data and Diagnostic Reasoning. *arXiv* **2025**, arXiv:2510.16899.
23. *ISO/TS Standard 15926-11:2023*; Simplified Industrial Usage of Reference Data Based on RDFS Methodology, 2nd, ed. International Organization for Standardization: Geneva, Switzerland, 2023. Available online: https://www.iso.org/obp/ui/es/#iso:std:iso:ts:15926:-11:ed-2:v1:en (accessed on 30 April 2026).
24. *ISO 15926-13:2018*; Industrial Automation Systems and Integration—Integration of Life-Cycle Data for Process Plants Including Oil and Gas Production Facilities. International Organization for Standardization (ISO): Geneva, Switzerland, 2018. Available online: https://www.iso.org/standard/70694.html (accessed on 28 June 2026).
25. Paslaru Bontas Simperl, E.; Tempich, C.; Sure, Y. ONTOCOM: A Cost Estimation Model for Ontology Engineering. In *The Semantic Web-ISWC 2006*; Cruz, I., Decker, S., Allemang, D., et al., Eds.; Lecture Notes in Computer Science; Springer: Berlin/Heidelberg, Germany, 2006; Volume 4273. https://doi.org/10.1007/11926078_45.
26. Babaei Giglou, H.; D'Souza, J.; Auer, S. LLMs4OL: Large Language Models for Ontology Learning. *arXiv* **2023**, arXiv:2307.16648.
27. LinkedIn Economic Graph Research Institute. Building LinkedIn's Skills Graph to Power a Skills-First World. LinkedIn. 2022. Available online: https://www.linkedin.com/blog/engineering/skills-graph/building-linkedin-s-skills-graph-to-power-a-skills-first-world (accessed on 20 April 2026).
28. Alsentzer, E.; Li, M.M.; Kobren, S.N.; et al. Few Shot Learning for Phenotype-Driven Diagnosis of Patients with Rare Genetic Diseases. *npj Digit. Med.* **2025**, *8*, 380. https://doi.org/10.1038/s41746-025-01749-1.
29. *ISO 8601-1:2019*; Date and Time—Representations for Information Interchange Part 1: Basic Rules. International Organization for Standardization (ISO): Geneva, Switzerland, 2019. Available online: https://www.iso.org/standard/70907.html (accessed on 28 June 2026).
30. Haug, A.; Zachariassen, F.; van Liempd, D. The Costs of Poor Data Quality. *J. Ind. Eng. Manag.* **2011**, *4*, 168–193. Available online: https://www.researchgate.net/publication/277237089_The_costs_of_poor_data_quality (accessed on 28 June 2026).
31. Simard, V.; Rönnqvist, M.; Lebel, L.; et al. A Method to Classify Data Quality for Decision Making Under Uncertainty. *ACM J. Data Inf. Qual.* **2023**. *15*, 17. https://doi.org/10.1145/3592534.

32. Vittor, L.; Arango, I.; Noori, A.; et al. A Scalable Platform to Build the Data Layer of Knowledge Graph AI. In Proceedings of the NeurIPS Workshop: New Perspectives in Graph Machine Learning, San Diego, CA, USA, 7 December 2025. Available online: https://openreview.net/forum?id=8OXD0yNZoD (accessed on 18 April 2026).
33. Wei, J.; Wang, X.; Schuurmans, D.; et al. Chain-of-Thought Prompting Elicits Reasoning in Large Language Models. In Proceedings of the Advances in Neural Information Processing Systems, New Orleans, LA, USA, 28 November–9 December 2022; Volume 35, pp. 5517–5528.
34. Dwork, C.; McSherry, F.; Nissim, K.; et al. Calibrating Noise to Sensitivity in Private Data Analysis. *J. Priv. Confidentiality* **2017**, *7*, 17–51. https://doi.org/10.29012/jpc.v7i3.405.
35. Rjoob, K.; McGurk, K.A.; Zheng, S.L.; et al. A multimodal vision knowledge graph of cardiovascular disease. *Nat. Cardiovasc. Res.* **2026***, 5*, 18–33. https://doi.org/10.1038/s44161-025-00757-4.
36. *LinkedIn Work Change Report 2025*; LinkedIn Economic Graph Research Institute: Sunnyvale, CA, USA, 2025.
37. Gartner. Gartner Predicts Over 40% of Agentic AI Projects Will Be Canceled by End of 2027. *Gartner Newsroom*, 25 June 2025. Available online: https://www.gartner.com/en/newsroom/press-releases/2025-06-25-gartner-predicts-over-40-percent-of-agentic-ai-projects-will-be-canceled-by-end-of-2027 (accessed on 18 April 2026).
38. Edjlali, R. Lack of AI-Ready Data Puts AI Projects at Risk. *Gartner Newsroom*, 26 February 2025. Available online: https://www.gartner.com/en/newsroom/press-releases/2025-02-26-lack-of-ai-ready-data-puts-ai-projects-at-risk (accessed on 31 March 2026).
39. The State of AI in the Enterprise. 2026 AI Report. Deloitte. 2026. Available online: https://www.deloitte.com/us/en/what-we-do/capabilities/applied-artificial-intelligence/content/state-of-ai-in-the-enterprise.html (accessed on 25 April 2026).
40. Gartner, Inc. Gartner Announces Top Predictions for Data and Analytics in 2026. *Gartner Newsroom*, 11 March 2026. Available online: https://www.gartner.com/en/newsroom/press-releases/2026-03-11-gartner-announces-top-predictions-for-data-and-analytics-in-2026 (accessed on 13 April 2026).
41. Zhu, X.; Li, Z.; Wang, X.; et al. Multi-Modal Knowledge Graph Construction and Application: A Survey. *IEEE Trans. Knowl. Data Eng.* **2024**, *36*, 715–735. https://doi.org/10.1109/TKDE.2023.3279574.
42. Kairouz, P.; McMahan, H.B.; Avent, B.; et al. Advances and Open Problems in Federated Learning. *Found. Trends Mach. Learn.* **2021**, *14*, 1–210.