ARTEFACT RESEARCH CENTER

A Knowledge Graph–Driven Enterprise Search System for Consulting Activities

Orel Mazor[A], Armita Khajeh Nassiri[A], Florence Benezit[A], Michael Völske[A] and Emmanuel Malherbe[A]
[A]Artefact Research Center, Paris, France

This paper presents a real-world consulting ontology and a corresponding Knowledge Graph (KG) designed to structure internal knowledge and capture learnings across diverse projects. Facing the challenge of leveraging unstructured material such as textual documents or commercial proposal slides, we apply LLM-based extraction and integrate the results with structured enterprise data through our ontology. Designed around internal usage patterns, this ontology enables the KG to populate a unified knowledge registry. This approach boosts retrieval efficiency and clarity by improving the information architecture, allowing teams to access and reuse enterprise knowledge more effectively. We first detail the system architecture and then discuss the operational challenges of data governance and user adoption.

1. Introduction

Consulting inherently relies on a wide range of domain knowledge spanning sectors, business topics, and technologies. Its core value proposition indeed lies in providing specialized expertise to clients lacking specific capabilities. In particular, the pure service business model generally implies that no Intellectual Property is retained internally, as ownership is purchased by the client [1]. However, consulting firms still possess a significant quantity of internal knowledge, such as the documents related to the commercial phase prior to project contractualization, or the key learnings and technologies consolidated at the end of a project.[1] Efficient knowledge sharing among consultants is essential for reusing and accumulating experience across teams and technologies. Artefact, a data and AI consulting firm with more than 1,700 employees worldwide, particularly faces this challenge due to its diversity of projects. Operating in 40 offices across 27 countries, it helps more than 1,000 organizations design and industrialize data-driven solutions tailored to their business needs.

When designing a knowledge-sharing system for a consulting firm, a practical challenge is that consultants spend most of their time on client delivery and have limited capacity to produce well-structured internal documentation. As a result, much of the firm's intellectual capital remains buried in unstructured materials with little metadata. Following these constraints, we developed an internal knowledge hub powered by a KG to support efficient knowledge exchange across the firm. This builds on recent work on knowledge graphs as a trusted foundation for enterprise question answering [4] and LLM-based knowledge graph construction [5], extending these ideas to the consulting domain where document-level extraction from slide decks poses distinct challenges. In this paper, we detail the system's design, implementation and practical impact.

2. Enterprise search for efficient knowledge sharing

The capabilities of LLMs and Retrieval-Augmented Generation (RAG) [3] have changed expectations around enterprise search. Consultants now expect systems that can answer contextual questions, find relevant experience, and point them to the person with the right expertise. However, achieving this level of expressiveness requires more than keyword-based indexing. It needs a shared semantic backbone that links core business concepts and contextualizes information extracted from unstructured documents.

To address this need, we first designed an ontology, based on user stories and frequently asked questions in internal channels (e.g., who has worked on a given technology or whether a previous mission produced relevant learnings). Building on this ontology and recent advances in language modeling, we developed a KG populated by an LLM extracting structured metadata from unstructured documents. This KG powers an enterprise knowledge platform where projects, experts, and technologies are connected, forming the semantic foundation for downstream applications.

[1]Our system strictly relies on authorized, anonymized content, independent of clients' projects or shared with explicit client permission for internal sharing.

Published at European Semantic Web Conference (ESWC) 2026
Artefact Research Center, Paris, France

2.1. Consulting Ontology

The consulting KG relies on a backbone of four interconnected blocks, a modular structure designed to enhance maintainability. Fig. 1 shows the main classes and properties of the Consulting Ontology.

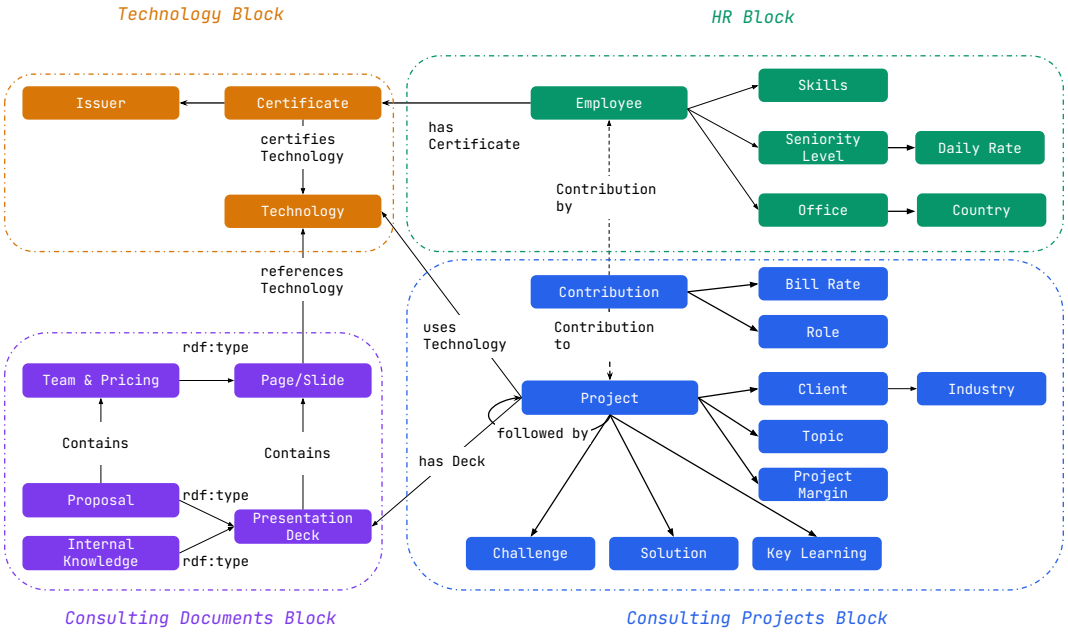


Figure 1 | Overview of the Consulting Ontology.

Consulting Projects Block. The `Project` class acts as the operational backbone of the graph. The `FollowedBy` property allows us to reconstruct complete mission lineage. The schema also captures the consulting problem-solution narrative via `Challenge`, `Solution`, and `Key Learning` class. A key design choice is the `Contribution` class, which links employees to projects. This intermediate node captures the specific `Role` and `Bill Rate` applied to a mission, enabling a clear financial distinction with the internal `Daily Rate` cost.

HR Block. Centered on the `Employee` class, this module manages staffing information and internal expertise mapping. As a design choice, its vocabulary aligns with the Human Resources Management Ontology to standardize employee-related concepts.[2]

Technology Block. This module standardizes the firm's technical landscape by modelling technologies as instances of the `Technology` class, reconciled with corresponding Wikidata items [2], together with a complementary `Certificate` class.

Consulting Documents Block. This module defines a taxonomy aligned with the consulting lifecycle (e.g., `Proposal`, `Internal Knowledge`). It also models document granularity at the `Slide` level, enabling the system to isolate and semantically classify important pages within larger documents.

2.2. System Architecture

In Figure 2, we define the system architecture overview, which includes four primary layers: data sources, data processing, knowledge management, and content delivery.

Data Sources. To populate the KG instantiated from the ontology in Fig. 1, we rely on two data sources: Structured Data from enterprise systems (HR, Finance, CRM, etc.) and Unstructured Data (slides, documents).

Data Processing. To instantiate the graph, we implement a hybrid pipeline that combines structured data extraction with LLM-based processing of unstructured documents. The workflow proceeds in 3 stages:

(1) Structured Data Extraction. Following the ontology schema, we define SQL mappings over enterprise

[2] https://oeg.fi.upm.es/index.php/en/ontologies/99-hrmontology/index.html

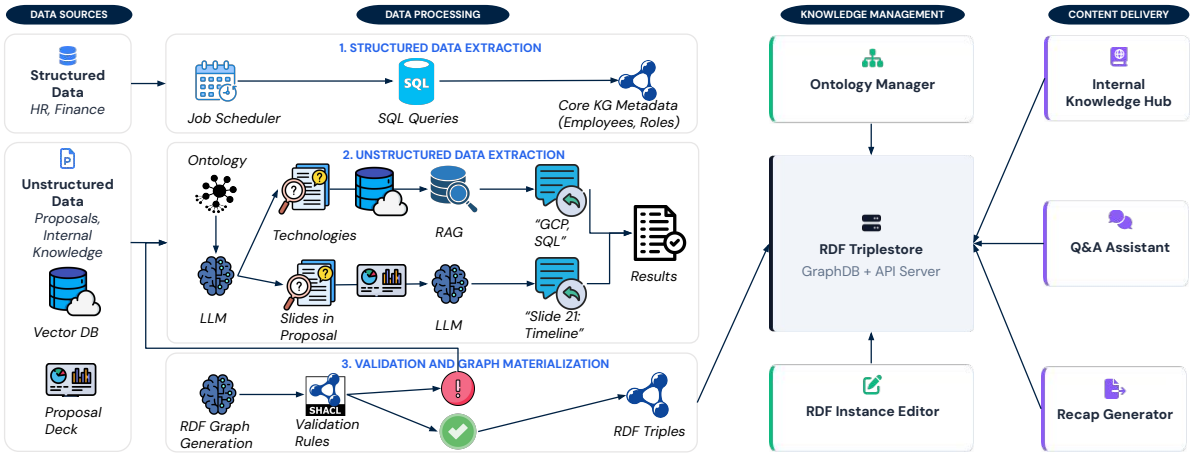


Figure 2 | Overview of the four-layer architecture used to build the consulting KG.

systems (HR, Finance, etc.) to populate the KG's structured backbone. These queries retrieve metadata such as employees, roles, etc. and run on a regular schedule to keep the graph synchronized and fresh with source systems.

(2) Unstructured Data Extraction We implemented an automated extraction pipeline to populate the ontology classes for each project. The LLM (Gemini-3-Pro) uses the ontology to guide extraction, identifying all information needed to instantiate the graph. For each class, we either apply a RAG approach over the presentations or directly prompt the model with relevant document context. Through this process, we extract structural elements such as `Challenge` and `Technology`, and classify high-value slides within presentations (e.g., Team and Pricing slides in a proposal).

(3) Validation and Graph Materialization. Extracted content is validated against SHACL constraints (e.g., at least 2 key learnings per project). Once these conditions are met, the LLM is prompted to generate RDF triples conforming to the ontology, and the resulting triples are inserted into the KG stored in our graph database (GraphDB[3]).

Knowledge Management. This layer forms the semantic backbone of the system, hosting the RDF triplestore and the API Server. It includes an *RDF Instance Editor*, which allows project owners to refine and update data, and an *Ontology Manager* used to govern vocabulary unification and synonym reconciliation. Operationally, a designated owner reviews and refines LLM-extracted metadata through the editor at the end of each project, a process that takes approximately 10 minutes.

Content Delivery. Leveraging the KG, we developed the *Internal Knowledge Hub* as a search interface for navigating firm credentials, alongside a Q&A Assistant for natural language querying. We also implemented a Project Recap Generator that formats project credentials into standardized slides for future proposals. All applications were deployed to a pilot group of consultants via a Streamlit web interface.

3. Evaluation and Lessons Learned

Evaluation. We assessed the system on two dimensions. For extraction quality, 10 project teams ingested their projects into the graph and modified less than 10% of the automatically extracted triples. For retrieval effectiveness, a comparative study with 30 consultants showed that the Knowledge Graph surfaced relevant insights in 80% of cases that were undiscoverable using the previous system.

User Adoption. Search utility alone proved insufficient to drive adoption: applications like the Project Recap Generator allowed consultants to perceive the system's immediate value and motivated their commitment to populate and refine the KG.

[3] https://graphwise.ai/components/graphdb/

References

[1] Carlsson, B., Dumitriu, M., Glass, J.T., Nard, C.A., Barrett, R.: Intellectual property (ip) management: organizational processes and structures, and the role of ip donations. The Journal of Technology Transfer **33**(6), 549–559 (2008)

[2] Lai, T., Ji, H., Zhai, C.: Improving candidate retrieval with entity profile generation for wikidata entity linking. In: Findings of the Association for Computational Linguistics: ACL 2022. pp. 3696–3711 (2022)

[3] Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Küttler, H., Lewis, M., Yih, W.t., Rocktäschel, T., Riedel, S., Kiela, D.: Retrieval-augmented generation for knowledge-intensive NLP tasks. In: Advances in Neural Information Processing Systems. vol. 33, pp. 9459–9474 (2020)

[4] Sequeda, J., Allemang, D., Jacob, B.: Knowledge graphs as a source of trust for LLM-powered enterprise question answering. Journal of Web Semantics **85** (2025)

[5] Zhu, Y., Wang, X., Chen, J., Qiao, S., Ou, Y., Yao, Y., Deng, S., Chen, H., Zhang, N.: LLMs for knowledge graph construction and reasoning: Recent capabilities and future opportunities. World Wide Web **27**(5), 58 (2024)