Ontology-based Knowledge Graph for Supply Chain and Operations Mapping*

João A.M. Santos[1,2], Rui M. Pinto[1], João M.C. Sousa[2], and Susana M. Vieira[2]

[1] Hovione Farmaciência S.A., Lumiar, Lisbon, Portugal
[2] IDMEC, Instituto Superior Técnico, Universidade de Lisboa, Lisbon, Portugal

Abstract. This paper introduces an ontology-based knowledge graph designed to enhance supply chain and operations mapping, built on the data from a company's enterprise resource planner (ERP). Comparative analysis with traditional data storage solutions reveals that the proposed approach significantly improves query efficiency, interpretability, and simplicity. Notably, the methodology expedites data access, particularly for complex queries, and utilizes simpler and more comprehensible queries. Furthermore, it advances the democratization of information within organizations, granting non-technical staff easier access to data.

Keywords: Knowledge Graph · Ontological Graph Database · Knowledge Representation · Database and Information Systems

1 Introduction

The 21st century has witnessed the transformation of supply chains (SCs) into intricate global networks. While essential to the world economy, these SCs remain susceptible to disruptions that expose challenges in transparency, resilience, and efficiency [7, 4]. Alongside technological advancements, companies must address issues of limited visibility, big data complexities, technical skill barriers, and trust surrounding data sharing [8, 1, 5]. Knowledge management concepts, such as knowledge graphs, offer promising solutions by organizing information systematically. Research demonstrates the potential of knowledge graphs in diverse SC applications, including power equipment, rail transit, and the semiconductor industry [13, 9, 11].

This paper proposes a knowledge graph with an ontological backbone for supply chain and operations mapping. The proposed methodology starts by defining the ontology that controls the knowledge graph and acts as a sort of meta-model of its workings. The methodology is compared with typical data storage solutions in terms of query efficiency, interpretability and simplicity. The proposed

* This work was supported by FCT, Fundação para a Ciência e a Tecnologia, I.P., under the PhD scholarship 2020.06065.BD (João A.M. Santos), Hovione Farmaciência S.A. (João A.M. Santos), and FCT, through IDMEC, under LAETA, project UIDB/50022/2020. The authors express their gratitude to the supporting institutions.

approach is built on the data of an ERP. The few publications that address the use of knowledge graphs on ERPs are much more high-level and focused on generating an overall graph of the entire enterprise [6, 3].

2 Ontology

An ontology can be defined as a structured way of representing knowledge within a specific area, including a taxonomy of classes, clear definitions of those classes, and the relationships between them [2].

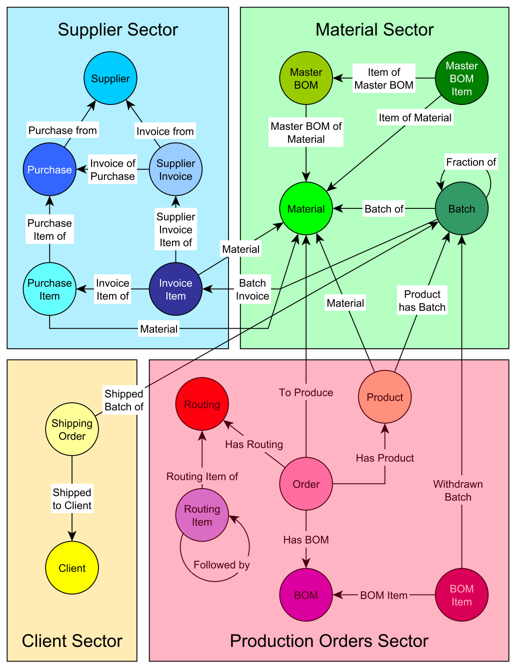

Fig. 1. Proposed ontology for the mapped graph database.

The proposed ontology shown in Figure 1 is divided into four sectors: suppliers, clients, materials and production orders. The suppliers sector deals mostly with the upstream supply chain; the clients sectors regards the downstream supply chain; the production orders sector regards the internal supply chain and operations; finally, the materials sectors addresses the materials resource planning. Each of these sectors tend to be mostly independent of each other, each featuring one or two cross-sector connections. Additionally, a few redundant connections were included to increase the degree of connectedness in some key nodes.

The main principles that were followed when creating the ontology were the following:

– **All classes are unique**

- **No ambiguous relations between two classes**
- **Instances respect classes and relations**
- **Attributes are consistent within instances of the same class**
- **All instances of classes have a unique ID**

The proposed ontology was built on data from the ERP, specifically modules of process planning and logistics. These are comprised of well defined relational tables, with an underlying schema, widely used by many companies all over the world. The ontology defines the connections between the data that are required for the purpose of SC and operations, enabling more direct connection between the data points, filtering out wrong data in the ETL process and summarizing the data into a more useful dataset.

3 Conclusions

The proposed methodology offers significant advantages over traditional relational databases accessed via SQL. Table 1 shows the results of 8 tests with increasing complexity comparing regular querying using SQL against the proposed methodology. The results show a significant reduction both in time and number of characters. This reduction becomes more pronounced with larger queries, showing reductions as high as 86% in time and 83% in length. Beyond efficiency, the methodology simplifies queries, making them more interpretable. This empowers non-technical workers to access data directly. In current paradigms, even basic data requests require technical workers proficient in SQL, creating bottlenecks and diverting skilled resources from more complex tasks. A well-structured ontology and an intuitive knowledge graph query language can bridge this gap, leading to the democratization of information within data-driven industries.

Query	Characters Reduction	Time Reduction
1	53.9%	315.0%
2	153.8%	8.2%
3	30.2%	78.9%
4	20.7%	49.2%
5	34.0%	15.7%
6	31.5%	2.9%
7	18.0%	30.2%
8	16.9%	14.4%

Table 1. Rate of reduction between relational databases and the proposed methodology for 8 tests of increasing complexity, regarding number of characters and time.

However, the methodology does have limitations. Data extraction and transformation inherently faces challenges like data conflicts and scalability concerns [10]. Additionally, the proposed ontology requires tailoring to the specific company's needs, data structure, and organization. It was designed with a focus on production planning and logistics, necessitating adaptation for other use cases.

A significant opportunity for improvement lies in training a large language model (LLM) on the ontology. This LLM could translate natural language requests into Neo4j queries, ultimately providing the desired information. This would remove the need for any programming knowledge, further democratizing data access. While ambitious, this approach is feasible due to the relative simplicity of Neo4j queries and has precedent in research using LLMs to generate SQL [12].

Bibliography

[1] Alpar, P., Schulz, M.: Self-service business intelligence. Business & Information Systems Engineering **58**, 151–155 (2016)

[2] Arp, R., Smith, B., Spear, A.D.: Building ontologies with basic formal ontology. Mit Press (2015)

[3] Blankenberg, C., Gebel-Sauer, B., Schubert, P.: Using a graph database for the ontology-based information integration of business objects from heterogenous business information systems. Procedia Computer Science **196**, 314–323 (2022)

[4] Brandon-Jones, E., Squire, B., Autry, C.W., Petersen, K.J.: A contingent resource-based perspective of supply chain resilience and robustness. Journal of Supply Chain Management **50**(3), 55–73 (2014)

[5] Byabazaire, J., O'Hare, G., Delaney, D.: Data quality and trust: Review of challenges and opportunities for data sharing in iot. Electronics **9**(12), 2083 (2020)

[6] Galkin, M., Auer, S., Vidal, M.E., Scerri, S.: Enterprise knowledge graphs: A semantic approach for knowledge management in the next generation of enterprise information systems. In: International Conference on Enterprise Information Systems, vol. 2, pp. 88–98, SCITEPRESS (2017)

[7] Kalaiarasan, R., Olhager, J., Agrawal, T.K., Wiktorsson, M.: The abcde of supply chain visibility: A systematic literature review and framework. International Journal of Production Economics **248**, 108464 (2022)

[8] Lennerholt, C., Van Laere, J., Söderström, E.: Implementation challenges of self service business intelligence: A literature review (2018)

[9] Li, S., Zhang, Y., Huang, M., Wu, H., Cai, W.: Building and using a supply chain knowledge graph applied to the rail transit industry. In: 2021 5th Asian Conference on Artificial Intelligence Technology (ACAIT), pp. 742–746, IEEE (2021)

[10] Ong, T.C., Kahn, M.G., Kwan, B.M., Yamashita, T., Brandt, E., Hosokawa, P., Uhrich, C., Schilling, L.M.: Dynamic-etl: a hybrid approach for health data extraction, transformation and loading. BMC medical informatics and decision making **17**(1), 1–12 (2017)

[11] Ramzy, N., Ehm, H., Durst, S., Wibmer, K., Bick, W.: Knowgraph-tt: Knowledge-graph-based transit time matching in semiconductor supply chains. INFOCOMMUNICATIONS JOURNAL **14**(1), 51–58 (2022)

[12] Song, Y., Wong, R.C.W., Zhao, X.: Speech-to-sql: toward speech-driven sql query generation from natural language question. The VLDB Journal pp. 1–23 (2024)

[13] Tang, Y., Liu, T., Shiy, J., Han, H., Liu, G., Dai, R., Wang, Z.: Ontology based knowledge modeling and extraction of power equipment supply chain. In: 2020 12th IEEE PES Asia-Pacific Power and Energy Engineering Conference (APPEEC), pp. 1–5, IEEE (2020)