Ontology-driven LLM framework for knowledge graph in smart buildings

Sebastien Gerin[1], Laurent Joblot[2], Nisrine Makhoul[1], Frederic Merienne[2]

[1]Smart, Sustainable, and Resilient Cities, ESTP, 11 Rue Sully, 21000 Dijon, France
[2]Arts et Métiers Institute of Technology, LISPEN / HESAM University / UBFC, Chalon-Sur-Saône, France

sgerin@estp.fr, laurent.joblot@ensam.eu nmakhoul@estp.fr frederic.merienne@ensam.eu

Abstract -

The rapid data growth in smart building environments requires advanced tools to integrate, interpret, and utilize this information effectively. Smart buildings generate vast and heterogeneous data streams, including sensor readings, occupancy metrics, and environmental conditions, which are critical for optimizing energy efficiency, enhancing occupant comfort, and enabling predictive maintenance. However, the lack of a structured approach to automatically connect and contextualize these data sources limits the insights that can be derived. To address these challenges, this paper presents a framework for assessing data in a knowledge graph by automatically retrieving entities and establishing relationships from diverse data sources, incorporating metadata standards and time-series data relevant to smart buildings. A standard ontology from the domain is used to drive the experiment, enabling the automatic construction of a semantic graph-based model for a real-world smart building environment. The framework's objective is to ensure information is comprehensible as a preliminary step to intelligent decision-making in data-driven smart buildings, enabling applications like fault detection, performance measurement, and energy auditing. The proposed approach explores the potential of large language models (LLMs) to automate data integration, reducing reliance on experts. This paper addresses existing literature gaps on metadata mapping and lays the groundwork for future advancements in digital twin technologies for smart building applications.

Keywords -

Ontology; Knowledge graph; Digital twins; data-driven approach; Large language models;

1 Introduction

Achieving global Net Zero Emissions (NZE) by 2050 is a critical goal for mitigating climate change, demanding transformative actions across various sectors. Among these, the Architecture, Engineering, Construction, and Operation (AECO) sector plays a pivotal role, necessitating an accelerated transition to align with the NZE scenario [1]. To this end, many countries have implemented stringent building energy codes and minimum performance standards aimed at reducing energy consumption and greenhouse gas emissions. Technological advancements and industry efforts have led to the rise of data-driven smart buildings (DDSBs), which integrate digitalization technologies to optimize energy use dynamically, enhance indoor environmental quality, and improve occupant experiences [2]. DDSBs generate vast amounts of data through Internet of Things (IoT) technologies, collecting real-time and historical information via advanced sensors. Effectively leveraging this data requires robust frameworks for systematic organization and analysis [2]. Complementing DDSBs, digital twins (DTs) further enhance operational efficiency by enabling predictive maintenance, informed decision-making, and cost minimization [3]. Researchers define a DT as "a virtual representation of a physical system (and its associated environment and processes) that is updated through the exchange of information between physical and virtual systems" [4]. Initially developed for aerospace applications, DTs now extend to diverse domains such as industrial manufacturing, healthcare, smart cities, education, and agriculture [5]. DTs are real-time digital replicas of physical entities or processes that enable a two-way flow of information. The virtual model can be updated in real-time, allowing simulations and informed decision-making to guide physical system adjustments [6]. When integrated with DDSBs, DTs leverage real-time and historical data to enhance data-driven processes, optimizing energy efficiency, indoor environmental quality, and overall system performance [2]. A critical aspect of DT implementation is the integration of a robust data schema, a structured framework essential for defining how data is organized, stored, and managed throughout the building's lifecycle [7]. Integrating appropriate data schemas requires specialized domain knowledge experts and a deep understanding of building systems. The lack of accessible, standardized solutions often hinders the practical deployment of DTs in the built environment, leading to underutilized data, ineffective decision-making, and inefficient practices with significant cost implications. This research introduces a framework that automates the mapping of data generated by smart buildings into a standardized model, serving as the foundational step in constructing a comprehensive DT. This paper seeks to address these limitations by answering the following questions:

1. How can we address the lack of generic tools for building digital twins in data-driven smart buildings?
2. What is the most effective method for automatically structuring heterogeneous data from a data-driven smart building into a coherent data schema?
3. Does our method perform well regardless of the data model used?

The paper is structured as follows: Section 2 reviews related work, providing context and highlighting existing approaches. Section 3 presents the proposed framework, which leverages Large Language Models (LLMs) for automated knowledge graph (KG) construction, significantly reducing the need for human intervention. The utilization of a prebuilt standard semantic graph ontology serves to enhance the organization of metadata, facilitating better efficiency and accuracy. Section 4 evaluates the framework through its implementation at the Dijon Metropolitan Campus, demonstrating its practical applicability in handling real-world datasets. Finally, Section 5 concludes the paper by summarizing the key findings and discussing potential avenues for future research and development.

2 Related work

Ontologies are essential for achieving data integration and interoperability in DDSBs and DTs. By providing a formalized vocabulary and structured taxonomy, they enable the systematic organization of domain-specific knowledge through well-defined classes, properties, and relationships [8]. Lygerakis et al. [9] identified over 40 ontologies relevant to energy optimization across the lifecycle of connected buildings. However, the authors highlighted that most remain theoretical, with limited real-world applications in operational settings. This literature review identifies several key ontologies as potentially relevant [10, 7], including the Building Ontology Topology (BOT) [11], SSN/SOSA [12], Brick [13], SAREF4BLDGG [14], the RealEstateCore ontology [15], and Google's Building Digital Ontology (DTO) [16]. While these ontologies provide a foundational schema for building data management, their implementation often necessitates significant domain expertise, posing challenges to scalability and automation. Knowledge graphs offer an alternative approach to managing heterogeneous and dynamic data in DDSBs and DTs. As graph-based data models, KGs structure knowledge in a machine-readable format, with nodes representing entities and labeled edges capturing relationships [17]. This flexibility enables KGs to augment digital twins by integrating real-time and historical data while supporting advanced analytics [18]. Unlike ontologies, KGs facilitate linking structured and unstructured data, offering actionable insights and enhancing decision-making capabilities. Despite these advantages, constructing KGs for DDSBs remains a non-trivial task, requiring the resolution of data ambiguities, alignment with specific application requirements, and managing diverse data sources [19]. Existing methodologies have primarily relied on semi-automated approaches. For instance, Chamari et al. [20] proposed a methodology to generate metadata schemas by correlating Building Automation System (BAS) identifiers with ontology classes through text search. This process, while partially automated, relies on human intervention for refining mappings. Similarly, Waterworth et al. [21] addressed inconsistencies in BAS metadata abbreviations using rules and language models to extract and classify semantic information. Koh et al. [22] developed a workflow utilizing machine learning algorithms to convert unstructured BAS metadata into structured formats like Brick. However, these approaches fall short of delivering fully automated solutions and often depend on significant manual effort or predefined rules.

To address these limitations, we propose a novel methodology for automating the construction of KGs from DDSB data. This approach proposes to limit the dependency on domain expertise by systematically mapping heterogeneous data into a standardized model, representing a foundational step toward the comprehensive implementation of digital twins. By directly addressing challenges such as data ambiguity and integration complexity, this methodology advances the field of DDSB data management and supports scalable, interoperable solutions. The following section details the design and implementation of our proposed framework.

3 The proposed approach

This study proposes a methodology to automate the generation of a KG from DDSB data using a large language model (LLM). As advanced generative AI systems, LLMs excel in processing and generating natural language, making them particularly effective for automating tasks such as entity and relationship extraction, competency question formulation, and metadata organization [23]. By reducing the dependence on manual effort and domain expertise, LLMs enhance both the scalability and precision of KG construction. However, a human-in-the-loop approach remains crucial for critical evaluation stages to ensure accuracy and reliability [24].

At the core of the methodology is the Retrieval-Augmented Generation (RAG) technique, further refined by the GraphRAG method[1]. RAG enhances the contextual relevance and accuracy of LLM outputs by combining semantic similarity retrieval from proprietary data sources with generative language capabilities [25]. This dual approach ensures that generated outputs align closely with domain-specific requirements, making RAG particularly valuable for knowledge-intensive domains such as DDSBs. GraphRAG extends RAG by introducing a KG-centric enhancement. The GraphRAG methodology organizes implicit knowledge from text into graph representation and creates communities based on the graph structure. The generated graph serves as context for RAG tasks, leveraging summaries from high-level communities to produce more diverse outputs. Comparative studies show that GraphRAG outperforms standard RAG methods in terms of accuracy, scalability, and cost-efficiency [26, 27].

Our approach is inspired by the GraphRAG process but uses a pre-built domain ontology as the context graph instead of generating one from raw data. In this manner, the domain ontology provides the context for the RAG operations, which empowers the retrieval process to extract entities from the domain-specific graph. This innovative approach addresses the challenges of data heterogeneity and complexity inherent to DDSBs, facilitating the creation of interoperable and standardized DTs. As illustrated in Figure 1, our proposed framework comprises three main phases: *data collection*, *KG construction*, and *KG visualization*. The following sections detail each phase of our process.

3.1 Data collection

The data collection phase represents the foundational step in our framework, aimed at systematically gathering and organizing the diverse data sources intrinsic to DDSBs. Smart buildings produce a wide spectrum of heterogeneous data, including IoT identifiers, time-series data from building management systems, and Building Information Modeling (BIM) files, often distributed across siloed and disparate databases. IoT metadata typically captures device-specific attributes such as type, operational parameters, control commands, and spatial topology, defined either by vendor-specific schemas or standardized frameworks. In our approach, the Building Operating System (BOS) plays a central role as the primary data aggregator, integrating information from interconnected devices and systems within the DDSB. This centralized integration ensures a unified flow of data, enhances compatibility with third-party applications, and facilitates the systematic identification and extraction of critical entities and relationships required for KG construction. By addressing the challenges of data heterogeneity and accessibility, this initial phase prepares the collected data for seamless transition to the subsequent phase.

[1] https://github.com/microsoft/graphrag

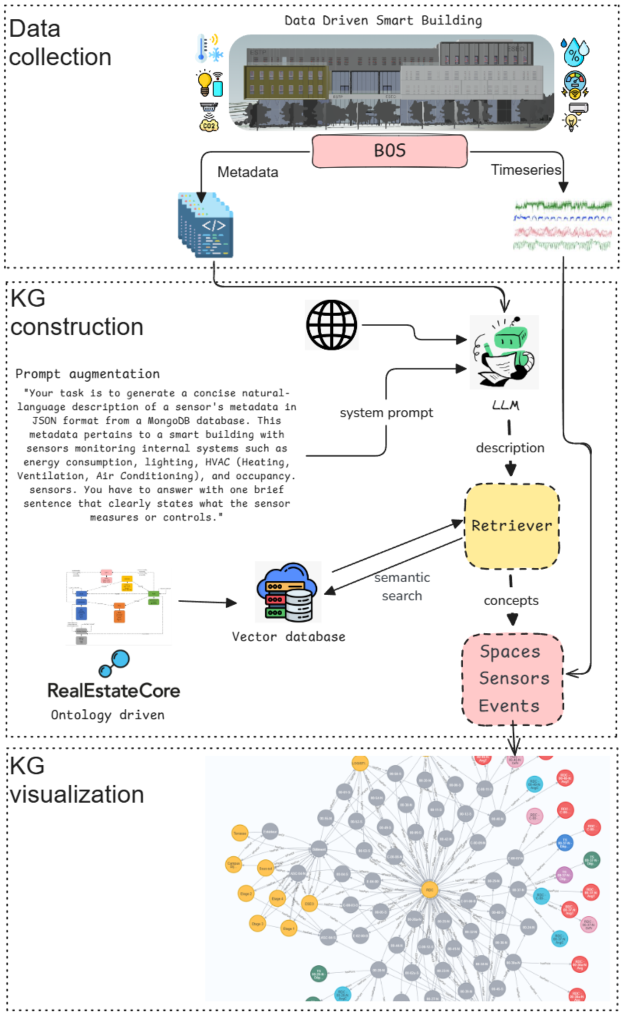

Figure 1. The architecture of the proposed approach.

3.2 KG construction

The second phase of our framework focuses on constructing a domain-specific KG by leveraging an enhanced LLM adapted to the GraphRAG methodology. This approach uses the LLM to generate concise natural-language descriptions from metadata provided by the DDSB and incorporates a predefined ontology to guide the semantic retrieval of appropriate concepts. The process begins with **(1) ontology integration**, where the GraphRAG system's context is set using a predefined ontology that semantically describes each asset of the DDSB, providing a structured framework for metadata interpretation. Next, **(2) metadata description generation** involves prompting the LLM to produce concise natural-language descriptions of the DDSB metadata, which includes dynamic data such as sensor readings (e.g., temperature, humidity, energy usage) and static data like spatial and topological information (e.g., classrooms, parking areas, laboratories, and building levels). Following this, **(3) concept retrieval** is performed, where a semantic similarity search is used to identify and retrieve the most relevant concepts from the ontology, aligning them with the metadata descriptions. Finally, **(4) KG construction** links the retrieved concepts with time-series data, organizing the information into a cohesive structure. Relationships are established between nodes representing topology, sensors, and time-series values, all formatted according to the ontology schema. This comprehensive process ensures that the constructed KG is semantically accurate, contextually relevant, and capable of representing the complex relationships inherent in DDSB data.

3.3 KG visualization

The final phase of our framework focuses on the visualization and evaluation of the constructed KG to ensure its accuracy and usability. The KG is exported to a graph database management system, which provides robust tools for querying and visualizing the graph's structure. Visualization plays a critical role in the validation process, enabling a thorough review of the interconnections between nodes. Specifically, this step involves verifying the alignment of the building topology, ensuring that sensors are correctly assigned to their respective locations, and confirming that observation values are accurately linked to the corresponding sensors, along with their associated timestamps. These visual checks are essential to identify and rectify any inconsistencies or errors that may arise during the KG construction phase. Additionally, the visualization phase facilitates an intuitive understanding of the underlying relationships and patterns in the data, making the KG more accessible to domain experts and stakeholders. This step concludes the methodology by ensuring that the KG is not only accurate and reliable but also ready for downstream applications, such as querying, analysis, and integration into digital twin systems.

4 Experimental results

This section presents the experimental results to evaluate the performance of the proposed method and analyze its effectiveness.

4.1 Evaluation setup

This study is conducted at the Dijon Metropolitan Campus[2], a $10,000\ m^2$ smart building certified under the Ready2Service (R2S) standard for its advanced digital infrastructure. The campus employs 14 types of sensors to monitor various parameters, including occupancy, temperature, CO2 levels, humidity, illuminance, and energy consumption, across 350 distinct areas encompassing 32 spatial categories such as waiting rooms and classrooms. The system stores the collected data in a schema-less MongoDB database in JSON format. A one-month subset of this time-series data amounts to approximately 11 GB and comprises over 50 million entries.

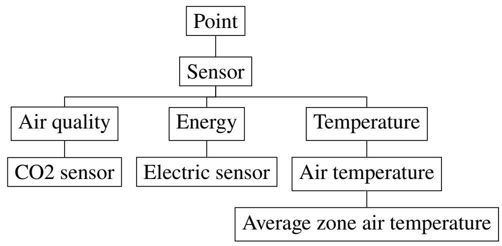

Figure 2. Part of REC ontology for sensor hierarchy.

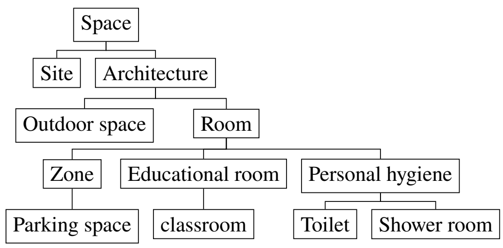

Figure 3. Part of REC ontology for space hierarchy.

To guide our methodology, we identify the RealEstateCore ontology (REC) as the most suitable framework due to its robust and comprehensive structure [15]. REC provides a detailed taxonomy that includes 240 sensor types (e.g., humidity, motion, temperature, lighting), as illustrated in Figure 2, and 92 spatial and architectural categories (e.g., back office, entrance, laboratory, warehouse), depicted in Figure 3. These categories are organized hierarchically to accommodate diverse use cases, enhancing the accuracy of KG construction and optimizing data retrieval processes. Moreover, REC employs the Label Property Graph (LPG) format, a data-centric model that facilitates real-time operations by enabling efficient storage and access of large datasets through graph databases [28]. Originally developed as an OWL ontology for RDF-based knowledge graphs, REC has been adapted into the Digital Twin Definition Language (DTDL)[3] to better align with user requirements and streamline deployment. DTDL, built on the JSON-LD format[4], ensures compatibility with the existing campus database, thereby eliminating the need

[2]https://www.estp.fr/en/dijon-campus

[3]https://learn.microsoft.com/en-us/azure/digital-twins/concepts-models

[4]https://json-ld.org/

for additional storage solutions while streamlining the integration process. This combination of hierarchical organization, real-time graph database performance, seamless integration, and compatibility with existing systems underscores REC's efficacy as the foundational framework for our approach.

4.2 Results and discussion

We present the results of applying the methodology to generate a KG at the Dijon Metropolitan Campus for a given day. The objective is to identify the pertinent classes within the REC ontology that align with the disparate points on the Dijon Metropolitan Campus and establish a relationship with the time series data. The resulting KG is exported to the Neo4J graph database[5] for storage and visualization. Figure 4 depicts a part of the resulting KG, in which the time-series data are linked to the standard semantics of the sensors and spaces of the Dijon Metropolitan Campus, encapsulated within the REC ontology model. Figure 5 illustrates an example of a node representing a cafeteria room that contains several attributes, including designation and display name, which have been retrieved from the metadata. This example demonstrates the hierarchy of the nodes, whereby the cafeteria room inherits from the food handling room, a room, and a space, corresponding to the REC ontology hierarchy and taxonomy. Figure 6 illustrates an example of a node representing a time series value connected to a node representing a temperature sensor. Node's properties include the temperature value and the timestamp. The objective is to demonstrate the enrichment of time-series data with standard semantics, facilitating its linkage to the sensors and equipment for efficient querying.

We evaluate the proposed methodology by presenting results from multiple perspectives, including processing execution time, repeatability, comparison of LLMs, concepts retrieval evaluation, and KG query evaluation.

4.2.1 Computational efficiency

The efficiency of the proposed methodology was assessed based on its execution time for generating the KG. The retriever identifies all relevant concepts within one minute, followed by KG construction in less than 20 seconds for a specific day's time-series data. Using a laptop equipped with an Intel Core i7 processor and 32 GB of RAM, the system generated $42,728$ JSON-LD documents (nodes) within this timeframe.

[5]https://neo4j.com/

4.2.2 Model consistency

The consistency of the retrieval methodology was assessed using three LLMs: Llama3.2[6], Mistral7B[7], and Gemma2[8]. Each model underwent 20 trials, with conservative parameter settings to minimize response variability. The results demonstrated that the retrieval process was reliable within individual models, with minimal fluctuations in performance due to the conservative configurations. However, notable discrepancies arose in how the models interpreted and retrieved concepts from the REC ontology. These differences highlight the distinct semantic inference strategies employed by each LLM, emphasizing the importance of carefully selecting the appropriate model for specific use cases and applications.

4.2.3 Model comparison and performance

The results shown in Table 1 illustrate pairwise comparisons that quantify the percentage of identical REC concepts retrieved by each LLM. Low values indicate that each LLM retrieved different concepts, while high values suggest greater homogeneity in the retrieved REC concepts across the models. The results generally show low to moderate values, indicating weak overlap in the retrieved concepts. This variation highlights differences in how each language model interprets and processes the semantics of the metadata and the REC ontology concepts, emphasizing the inherent differences in their inference mechanisms.

Table 1. Retrievability analysis across LLMs for 14 sensors (Se.) and 32 spaces (Sp.).

LLM	Llama		Mistral		Gemma	
	Se.	Sp.	Se.	Sp.	Se.	Sp.
Llama	-	-	28%	53%	21%	68%
Mistral	28%	53%	-	-	28%	50%
Gemma	21%	68%	28%	50%	-	-

4.2.4 Semantic retrieval accuracy

A manual qualitative assessment was conducted to evaluate the identification of 32 spaces and 14 sensor types from the REC ontology. Each retrieved concept was categorized as "correct," "ambiguous," or "incorrect." The ground truth for each sensor and space category was established through direct analysis of the data or reference to the BIM model. For instance, an incorrect retrieval was observed when the Gemma model classified a *meeting room* as a *workshop room*, while both the Mistral and Llama models correctly identified a *conference room*. In cases where ambiguity was present, all LLM models

[6]https://www.llama.com/
[7]https://docs.mistral.ai/
[8]https://github.com/google-deepmind/gemma

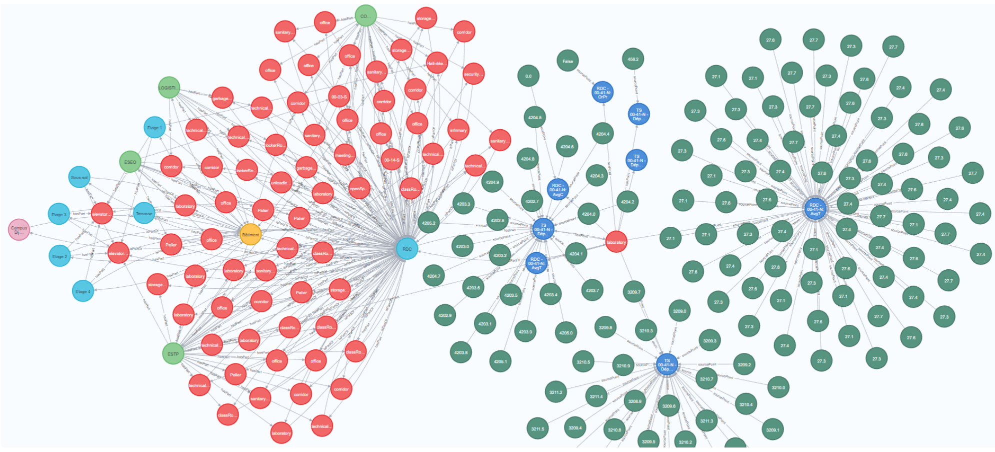

Figure 4. Part of the generated KG for the Dijon Metropolitan Campus: red nodes represent spaces, blue nodes represent sensors, and green nodes represent values.

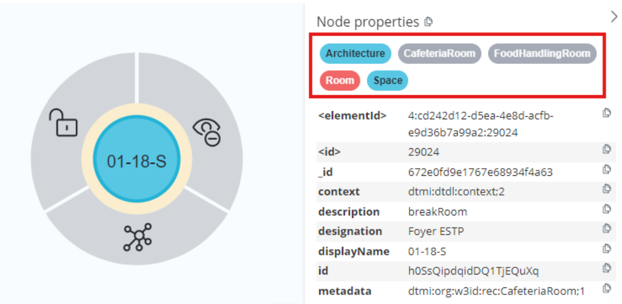

Figure 5. Cafeteria room node with hierarchy.

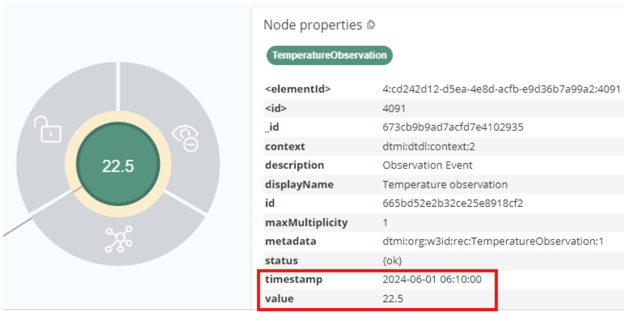

Figure 6. Temperature observation node with values.

classified *toilets* as *personal hygiene rooms*, despite the explicit existence of the toilet concept in the REC ontology. An example of a correct retrieval is the occupancy sensor, which was accurately mapped to the corresponding concept within the REC ontology. The Dijon Metropolitan Campus employs energy sensors that monitor electricity consumption in kilowatt-hours (kWh) across various equipment, such as lighting, power outlets, and heating systems. The most appropriate concept for this data in the REC ontology is the *Electric Energy Sensor*. However, depending on the LLM, the retrieved concepts varied between Energy Usage Sensor and Energy Sensor. While both of these concepts are valid, they do not fully align with the ground truth, and as such, these responses were categorized as "ambiguous," leading to a loss of semantic precision regarding the type of electrical energy. Table 2 summarizes the performance of the models, revealing that the Llama model consistently outperformed the others, exhibiting high accuracy with minimal errors in both space and sensor retrieval tasks. The Gemma model demonstrated moderate performance, excelling in space retrieval but encountering more ambiguities in sensor retrieval. The Mistral model, in contrast, showed the weakest performance, particularly in space retrieval, with a significant number of incorrect responses.
However, ambiguous responses are not necessarily erroneous, though they do not invariably correspond with the ground truth. The models performed consistently across various scenarios, achieving an overall correct identification rate of 73% when ambiguous answers were considered accurate

Table 2. Retrievability evaluation of sensors (Se.) and spaces (Sp.) across LLMs: distribution of correct, ambiguous, and incorrect retrievals.

LLM	Correct		Ambiguous		Incorrect	
	Se.	Sp.	Se.	Sp.	Se.	Sp.
Llama	12	27	2	2	0	3
Mistral	5	17	3	2	6	13
Gemma	5	20	8	2	1	10

4.2.5 KG query evaluation

We leverage the semantics of metadata and time-series data structured within the REC ontology to assess the generated KG. The relevance of semantic relationships is explored through a series of Cypher queries[9], a declarative query language designed for use with the Neo4j graph database. The evaluation involved conducting automated queries on the KG and validating the results against a predefined ground truth. A total of 15 queries were formulated and categorized into three distinct types: **(1) room type query** counting, which examines the distribution of room types within the KG; **(2) sensor type query** counting, focusing on the frequency of occurrence of various sensor types; and **(3) sensor-to-room relationships**, aimed at evaluating the connections between sensors, their associated time-series data, and the specific spaces they monitor. For a complete list of queries, please refer to the file `queries_for_kg_evaluation.md` in the GitHub repository[10]. The results of these evaluations are summarized in Table 3, which presents the number of correct answers for each LLM. The ground truth was established by querying the original MongoDB database or through manual analysis of data from the BIM model. The Llama model demonstrated the most consistent performance, achieving near-perfect scores across all query types, thereby illustrating its robust ability to handle both sensor and spatial data. The Mistral model performed reasonably well in sensor-to-room mapping but exhibited weaker results in the room type and sensor type queries. The Gemma model, while excelling in sensor-related queries, faced challenges in accurately identifying room types. These findings highlight the effectiveness of using automated unit tests for evaluating and refining the performance of knowledge graph queries.

Table 3. Automatic query evaluation.

LLM	Room type	Sensor type	Sensor-to-room
Llama	4/5	5/5	4/5
Mistral	2/5	3/5	4/5
Gemma	1/5	5/5	5/5

5 Conclusion and future work

This paper presents an LLM-based framework that leverages an ontology to guide the automatic construction of a KG, facilitating the management of data in smart buildings. By integrating LLMs, the framework reduces reliance on human expertise while improving the accuracy of KG development. Experimental results demonstrate the effectiveness of the REC ontology in guiding KG creation from unstructured data. However, while the method accelerates the process, human verification is still required to ensure the correctness of identified concepts, introducing potential delays and biases. These preliminary findings suggest promising directions for future research, including integrating BIM assets into the KG [29] to enhance organization and interoperability. Expanding the REC ontology to address ambiguous or incorrect concepts could further refine the KG's semantic alignment. Additionally, exploring the use of alternative ontologies, such as Brick, would increase user autonomy and broaden the framework's applicability. Further testing with metadata from other smart buildings would also be beneficial. Once established, the KG can serve as a foundation for LLMs, enabling natural language queries that improve accessibility for diverse stakeholders. Moreover, applying advanced graph algorithms, such as graph neural networks, could offer valuable insights, supporting classification and clustering tasks that enhance decision-making in complex smart building scenarios.

6 Acknowledgement

We thank Primatice, a subsidiary of Vinci Facilities, for their financial and material support. Primatice is developing TwinOps[11], a platform for managing digital twins of real estate assets to improve smart building operations and maintenance.

References

[1] IEA. IEA (2022), all countries targeted for zero-carbon-ready codes for new buildings by 2030, iea, paris, Accessed: 01/10/2024.

[2] J. Candanedo, D. Blum, Z. Chen, G. Fierro, V. Gori, H. Johra, H. Madsen, A. Marszal-Pomianowska, Z. O'Neill, O. Pradhan, et al. Data-driven smart buildings: State-of-the-art review. *Energy in Building and Communities Programme*, page 103, 2023.

[3] Y. Mousavi, Z. Gharineiat, A. A. Karimi, and S. Gonizzi Barsanti. Digital twin technology in built environment: A review of applications, capabilities and challenges. *Smart Cities*, 7(5):2594–2615, 2024.

[4] E. VanDerHorn and S. Mahadevan. Digital twin: Generalization, characterization and implementation. *Decision support systems*, 145:113524, 2021.

[5] M Mateev. Industry 4.0 and the digital twin for building industry. *Industry 4.0*, 5(1):29–32, 2020.

[9] https://neo4j.com/docs/cypher-manual/current/introduction/
[10] https://tinyurl.com/3heep8pu
[11] https://www.twinops.com/

[6] M. Attaran and B. G. Celik. Digital twin: Benefits, use cases, challenges, and opportunities. *Decision Analytics Journal*, 6, 2023.

[7] G. Fierro and P. Pauwels. Survey of metadata schemas for datadriven smart buildings (annex 81). *Energy in Buildings and Communities Technology Collaboration Programme*, 2022.

[8] T. R. Gruber. A translation approach to portable ontology specifications. *Knowledge acquisition*, 5 (2):199–220, 1993.

[9] F. Lygerakis, N. Kampelis, and D. Kolokotsa. Knowledge graphs' ontologies and applications for energy efficiency in buildings: A review. *Energies*, 15(20): 7520, 2022.

[10] Z. Qiang, S. Hands, K. Taylor, S. Sethuvenkatraman, D. Hugo, P. G. Omran, M. Perera, and A. Haller. A systematic comparison and evaluation of building ontologies for deploying data-driven analytics in smart buildings. *Energy and Buildings*, 292, 2023.

[11] M. H. Rasmussen, M. Lefrançois, G. F. Schneider, and P. Pauwels. Bot: The building topology ontology of the w3c linked building data group. *Semantic Web*, 12(1):143–161, 2021.

[12] M. Compton, P. Barnaghi, L. Bermudez, R. Garcia-Castro, et al. The ssn ontology of the w3c semantic sensor network incubator group. *Journal of Web Semantics*, 17:25–32, 2012.

[13] B. Balaji, A. Bhattacharya, G. Fierro, J. Gao, J. Gluck, et al. Brick: Metadata schema for portable smart building applications. *Applied energy*, 226: 1273–1292, 2018.

[14] M. Poveda-Villalón and R. García-Castro. Extending the saref ontology for building devices and topology. In *LDAC*, pages 16–23, 2018.

[15] K. Hammar, E. O. Wallin, P. Karlberg, and D. Hälleberg. The realestatecore ontology. In *Proceedings of ISWC, Auckland, New Zealand, October 26–30, 2019*, pages 130–145. Springer, 2019.

[16] K. Berkoben, C. El Kaed, and T. Sodorff. A digital buildings ontology for google's real estate. In *ISWC (Demos/Industry)*, pages 392–394, 2020.

[17] L. Ehrlinger and W. Wöß. Towards a definition of knowledge graphs. *SEMANTiCS (Posters, Demos, SuCCESS)*, 48(1–4):2, 2016.

[18] N. Sahlab, S. Kamm, et al. Knowledge graphs as enhancers of intelligent digital twins. In *Proceedings of the 4th IEEE ICPS*, pages 19–24, 2021.

[19] V. Chaudhri, C. Baru, et al. Knowledge graphs: introduction, history and, perspectives. *AI Magazine*, 43(1):17–29, 2022.

[20] L. Chamari, J. van der Weijden, L. Boonstra, S. Hoekstra, E. Petrova, and P. Pauwels. Metadata schema generation for data-driven smart buildings. In *Proceedings of LDAC*, pages 139–150. CEUR-WS. org, 2023.

[21] D. Waterworth, S. Sethuvenkatraman, and Q. Z. Sheng. Building metadata inference using a transducer based language model. *arXiv preprint arXiv:2212.01964*, 2022.

[22] J. Koh, B. Balaji, D. Sengupta, J. McAuley, R. Gupta, and Y. Agarwal. Scrabble: transferrable semi-automated semantic metadata normalization using intermediate representation. In *Proceedings of the 5th Conference on Systems for Built Environments*, pages 11–20, 2018.

[23] T. Brown, B. Mann, N. Ryder, M. Subbiah, J. D. Kaplan, P. Dhariwal, et al. Language models are few-shot learners. *Advances in neural information processing systems*, 33:1877–1901, 2020.

[24] V. K. Kommineni, B. König-Ries, and S. Samuel. From human experts to machines: An llm supported approach to ontology and knowledge graph construction. *arXiv preprint arXiv:2403.08345*, 2024.

[25] P. Lewis, E. Perez, A. Piktus, F. Petroni, V. Karpukhin, N. Goyal, et al. Retrieval-augmented generation for knowledge-intensive nlp tasks. *Advances in Neural Information Processing Systems*, 33:9459–9474, 2020.

[26] D. Mozolevskyi and W. AlShikh. Comparative analysis of retrieval systems in the real world. *arXiv preprint arXiv:2405.02048*, 2024.

[27] D. Edge, H. Trinh, N. Cheng, J. Bradley, A. Chao, A. Mody, S. Truitt, and J. Larson. From local to global: A graph rag approach to query-focused summarization. *arXiv preprint arXiv:2404.16130*, 2024.

[28] R. Angles. The property graph database model. In *Proceedings of the Workshop on Foundations of Data Management*, 2018.

[29] C. Ramonell, R. Chacón, and H. Posada. Knowledge graph-based data integration system for digital twins of built assets. *Automation in Construction*, 156: 105109, 2023.