Proceedings of International Structural Engineering and Construction, 12(1), 2025
Integrating Technology for Advanced Practices in Structural Engineering and Construction
Edited by Vimonsatit V., Askarinejad H., Pour M., Singh, A., and Yazdani S.

www.doi.org/10.14455/ISEC.2025.12(1).CON-08

GRAPHRAG-BASED DECISION SUPPORT SYSTEM INTEGRATING AND SELECTING SUSTAINABLE BUILDING PRODUCTS

DOMINIK STELLMACHER[1], KARINA GROßE LÖGTEN[2], SEBASTIAN THEIßEN[1,2], AGNES KELM[1], and ANICA MEINS-BECKER[1]

[1]*Dept of Digital Design, Construction and Operation, Univ of Wuppertal, Wuppertal, Germany*
[2]*List Eco, Cologne, Germany*

In light of rising demands for ecological, economic, and social sustainability in the construction sector, data-driven decision support is now essential. This article introduces a GraphRAG (Graph-based Retrieval-Augmented Generation)-based pipeline that leverages Large Language Models (LLMs) to extract and structure unstructured product information from datasheets and Environmental Product Declarations (EPDs). We then map this data into a Neo4j property graph. A custom in-house Retrieval-Augmented Generation (RAG) framework supports context-aware queries and analyses of key sustainability indicators for building materials. Through an intuitive user interface, engineers, architects, and project owners can quickly integrate economic data - such as detailed cost evaluations - alongside technical and environmental metrics. This approach widens the basis for decision making. The pipeline increases transparency in material selection and life cycle assessment. It enables more informed insights into resource efficiency and environmental impacts. Our results demonstrate that combining LLMs with a knowledge graph environment improves data quality and decision making by delivering context-relevant information in real time. Overall, the GraphRAG pipeline offers a novel contribution to data-centric sustainability assessment in the built environment and provides a robust foundation for future-proof material decisions.

Keywords: Large language models (LLMs), Knowledge graph, Automated information extraction, Chain-of-thought prompting, Multi-criteria evaluation.

1 INTRODUCTION

Sustainable construction requires evaluating materials on multiple criteria - technical performance, environmental impact, economic feasibility, and socio-cultural factors (Alkasar and Yahya 2023). Traditionally, multi-criteria decision-making (MCDM) methods have aided material selection (Alam Bhuiyan and Hammad 2023), though these often rely on manually compiled data and are time-consuming. With Large Language Models (LLMs), it is now possible to automatically extract and structure knowledge from unstructured documents (e.g., product data sheets and EPDs (Schneider-Marin *et al.* 2022). GraphRAG (Graph-based Retrieval-Augmented Generation) leverages LLMs to create a knowledge graph from text corpora, enabling more structured retrieval than pure text search (Zhang *et al.* 2025). This paper presents a GraphRAG-based decision support system for sustainable construction products that automatically classifies products, extracts attributes into a knowledge graph, and supports interactive queries across sustainability dimensions. It demonstrates how chain-of-thought prompting and domain-specific extraction rules enhance the quality of information extraction. The following sections describe the state of the art, methodology, and evaluation using exemplary construction products.

2 STATE OF THE ART

Current literature highlights LLMs' potential for automated document analysis in construction informatics, with sustainability frameworks integrating EPD data into life cycle analysis for product selection. Meanwhile, retrieval-augmented generation (RAG) effectively grounds LLM outputs by retrieving relevant text segments (Lewis *et al.* 2020), yet Larson and Truitt (2024) note that conventional RAG systems often lack multi-step reasoning across disparate data sources.

Knowledge graphs (KGs) have emerged as a powerful tool to enhance LLM performance in question-answering systems. The GraphRAG approach from Microsoft Research (Larson and Truitt 2024) utilizes LLM-generated knowledge graphs to better contextualize queries and improve response explainability. Hou *et al.* (2024) demonstrated that integrating knowledge graphs with LLMs can markedly improve retrieval performance and response transparency. The present study builds on these insights by focusing on a GraphRAG-based approach for the selection of sustainable construction products.

3 METHOD

The developed decision support system implements a GraphRAG pipeline that ingests product documents and constructs a structured knowledge graph of construction products. The core concept is to leverage LLMs for both understanding (i.e., classification) and information extraction - supported by domain-specific knowledge. The pipeline consists of seven steps (as illustrated in Figure 1), with a primary focus on Step 2: the classification and extraction of product information.

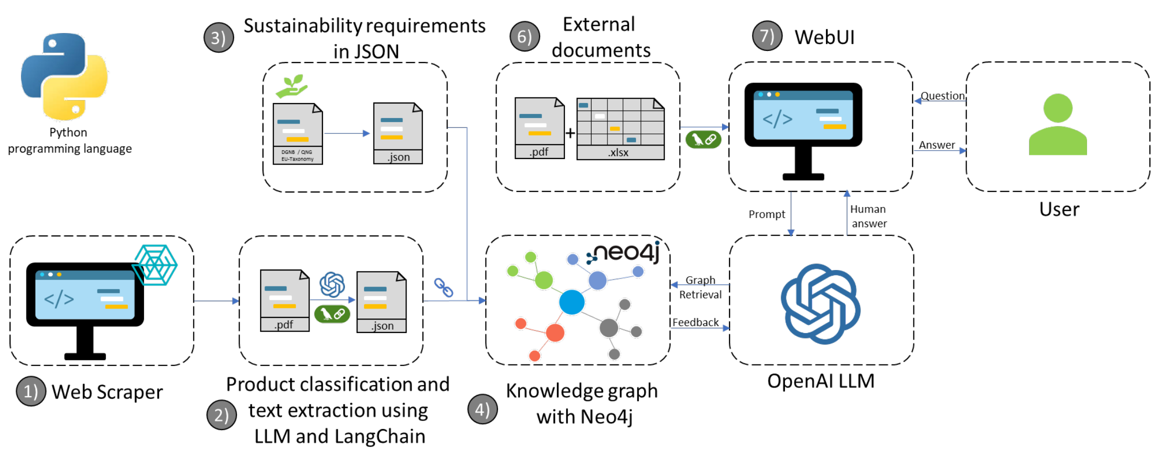

Figure 1. Systematic Overview of the GraphRAG Pipeline (own illustration).

In Step 1, product data sheets are collected using web scrapers that traverse manufacturer websites' HTML structures, identifying and structurally storing documents based on file extensions (e.g., .pdf). It assumes that the documents are original digital PDFs; scanned PDFs are excluded as there is no OCR integration. Step 2 then executes the following four sub-steps:

- **Product Type Classification using LLMs & Embeddings**

Product documents are first projected into a vector space via an OpenAI embedding model. The resulting vectors are compared with reference embeddings from the eClass classification system's product type definitions. The vector with the highest similarity suggests a product type - for example, one data sheet yielded "lime-sand brick" with a similarity score of 0.6189. Then, OpenAI's GPT-4-based model (internal 'GPT 4o-mini') validates this suggestion using a Chain-of-Thought (CoT) prompt (Wei *et al.* 2022) by analyzing the document's content and establishing broader product groups (e.g., masonry units) and trades (e.g. masonry work). For the lime-sand brick, the analysis is as follows:

(i) **Keyword Identification:** The document cites "plan stone," compressive strength class 28, a thermal conductivity of 1.8 W/(mK), and notes its use in load-bearing interior walls.
(ii) **Product Type Comparison:** Lime-sand bricks - as manufactured masonry units – exhibit these characteristics, unlike alternatives such as natural stone, brick, or concrete.
(iii) **Final Classification:** This leads to the definitive assignment as "lime-sand brick".

- **Variant Identification using Chain-of-Thought**

Many data sheets describe not a single product but a family of related variants that differ, for example, in dimensions. A specialized CoT prompt is employed to systematically capture these variations. The LLM incrementally analyzes the document, extracting key features such as dimensions (L×W×H) and dependent attributes like thermal conductivity and sound insulation. The result is output as a JSON array, where each variant is represented as an object with its essential properties. For instance, variants such as "Product A 248x175x248" or "Product A 248x200x248" are extracted, along with their corresponding dimensions, thermal conductivity, and sound insulation values, enabling precise differentiation in later queries.

- **Product Type-Specific Attribute Extraction**

After determining the product type and variant, a specialized extraction process applies custom rules to extract only pertinent attributes - such as compressive strength, bulk density, strength class, thermal conductivity, and dimensions for masonry units or drying time, density, and minimum application thickness for adhesives. Different terminologies (e.g., thermal transmittance versus U-value) are standardized. In addition to technical data, ecological and socio-cultural information are extracted: socio-cultural attributes (e.g., VOC content from safety data sheets) similarly to technical properties, and ecological information from EPDs is parsed using Python libraries (e.g., re and openpyxl) based on a standardized structure. All extracted data are consolidated into a structured JSON format and stored as separate nodes in the knowledge graph (Step 4).

- **Validation Loop with Confidence Score**

To enhance reliability, a validation loop re-evaluates the original text excerpts and the extracted value-unit pairs. The LLM confirms whether the values are correct ("OK") or provides corrected entries in the format "value|unit". This process is repeated until two consecutive identical results are achieved or a predefined limit is reached. Fields that are consistently validated receive a confidence score of 1.0, while uncertain values are assigned a lower score (0.5). All validation outcomes are logged, and the confirmed attributes are stored in the knowledge graph, ensuring precise and traceable queries.

After product information is stored, Step 3 links it with sustainability requirements. Criteria from the DGNB Systems (2024) and EU Taxonomy (Climate Positive Europe Alliance 2023) catalogs are converted into JSON, and the extracted values (e.g., the VOC content of an adhesive) are compared against these criteria. The resultant relationship - labeled "FULFILLS_REQUIREMENTS_OF" - is recorded in Neo4j, the central knowledge graph of the GraphRAG pipeline, to enable targeted queries for products that meet specific sustainability standards. In Step 4, the extracted data for each construction product are organized into the categories of technical, ecological, economic, and socio-cultural. Finally, the WebUI developed in Step 5 serves as the interface between the user and the GraphRAG pipeline, facilitating precise and transparent queries based on the knowledge stored in Neo4j.

4 RESULTS

To validate and evaluate the decision support system, both quantitative and qualitative analyses were conducted, with the following findings:

4.1 Quantitative Analysis of Extraction Performance

The system was tested on eight construction products, including two dispersion adhesives (DA), a lime-sand brick (LSB), a cellular concrete block (ACB), a multi-mortar (MM), a gypsum board (PB), a rafter insulation (RI), and a concrete (C). Accuracy was assessed by comparing extracted values with the original documents. Table 1 summarizes the results, detailing the correct product type, the embedding model's suggested type (with its mathematical similarity), the final CoT, the number of variants, the count of extracted attributes (categorized as technical, ecological, economic, and socio-cultural), the correctness rate, the confidence score (CS) for technical/social fields, the average extraction duration per variant, and the document types used (e.g., technical data sheet (PDS), EPD, and safety data sheet (SDS)).

Table 1. Extraction results by product type.

Product Type	Embedding Suggestion	Final CoT Classification	Number of Variants	Number of Attributes	Correctness Rate [%]	CS	Documents
LSB	LSB (0.6189)	LSB	3	9/263/0/0	100	1.0	PDS, EPD
ACB	PB (0.5479)	ACB	6	7/219/0/0	99[1]	0.9	PDS, EPD
MM	MM (0.6004)	MM	1	8/173/0/3	100	0.9	PDS, EPD, SDS
DA	PB (0.4570)	DA	1	6/173/0/4	100	0.97	PDS, EPD, SDS
DA	DA (0.6308)	DA	1	6/0/0/4	100	1.0	PDS, SDS
PB	PB (0.6141)	PB	5	5/265/0/0	100	1.0	PDS; EPD
RI	DA	RI	11	9/150/0/0	100	1.0	PDS, EPD
C	MM	C	1	6/0/0/3	100	1.0	PDS, SDS

High ecological attribute counts result from automated capture of numerous environmental indicators and resource consumption figures - as defined by DIN EN 15804 across life cycle phases (e.g., A1–A3, B1, etc.) - plus metadata from title pages (e.g., validity and declared unit). The final CoT classification accurately reflected actual product properties, even though the embedding model initially misclassified four out of eight products - errors that were reliably corrected through CoT analysis. Overall, the correctness rate was 99% or higher - with one minor error in a cellular concrete block variant (incorrect strength value). Most products achieved a confidence score of 1.0, with a few (e.g., the multi-mortar product) at 0.9, and average extraction durations ranged from 29 to 79 seconds per variant. The system demonstrates exceptionally high extraction and classification quality, even with extensive EPD data.

[1] 5 variants → No error | 1 variant → Strength error: Expected "6", extracted "F90-A, REI A180, EI 180".

4.2 Qualitative Evaluation of Query Responses

Figure 2 illustrates communication via the WebUI (corresponding to Step 5 in the pipeline shown in Figure 1), demonstrating how the system precisely and concisely answers queries from the four sustainability categories – technical, ecological, economic, and socio-cultural.

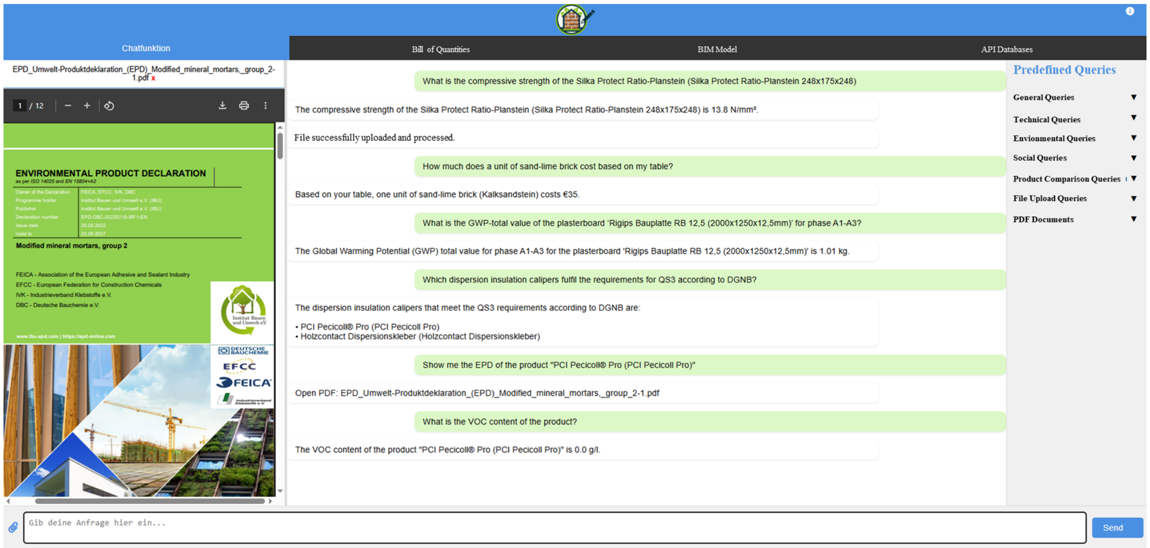

Figure 2. WebUI as the interface between the user and the GraphRAG pipeline (own illustration).

The illustrated communication via the WebUI (Step 5 in Figure 1) demonstrates how the system precisely answers queries from the four sustainability categories. For each category, a user query is processed by the LLM using the structured knowledge graph and integrated PDF viewer, with external documents (e.g., price lists) uploaded for economic information. The system was queried in all four dimensions; examples of its precise, context-aware answers are as follows:

- **Technical queries** (e.g., on compressive strength or performance metrics such as 13.8 N/mm² for Product A) are answered exactly using data from the sheets.
- **Ecological queries** yield precise answers through the integration of extensive EPD data, even after a formatting error (omission of "CO_2-equivalent") was detected and corrected.
- **Economic queries** produce accurate responses thanks to external cost information (e.g., pricing from uploaded tables).
- **Socio-cultural** queries offer well-founded recommendations based on normative and health-related data.

The provision of PDF documents via the PDF viewer—whose filenames are stored in the knowledge graph and retrievable as needed—ensures control and access to additional information, enhancing both transparency and trust in the automated responses. Additionally, the ReAct framework (Yao *et al.* 2023) with memory functionality facilitates smooth communication by caching previous inputs for use in subsequent queries. Furthermore, the scalability of the approach is assessed positively: by leveraging the Neo4j knowledge graph for data storage and the simplicity of query operations, the system is expected to handle substantially larger document volumes with ease. Overall, the evaluation shows that the LLM consistently delivers accurate information based on the extracted data, effectively supporting users in selecting sustainable construction products.

5 CONCLUSIONS

This work introduces the development of a decision support system that leverages LLM-based knowledge graph construction (GraphRAG) for the automated classification and attribute extraction of sustainable construction products. Despite extraction errors - such as those encountered with the cellular concrete block - more than 99 % of the attributes are accurately captured on average, providing precise, contextually interpreted responses and well-founded material decisions. The results demonstrate the potential of AI methods to overcome data silos in construction and establish structured decision-making foundations. The pipeline is currently being further refined to address remaining errors before its release as an open-source tool.

Acknowledgements

The presented results originate from the research project "NaConBau" – funded by the German Federal Ministry of Research, Technology and Space as part of the DATIpilot funding program.

References

Alam Bhuiyan, M. M., and Hammad, A., *A Hybrid Multi-Criteria Decision Support System for Selecting the Most Sustainable Structural Material for a Multistory Building Construction*, Sustainability, MDPI, 15(4), 3128, February 2023.

Alkasar, B. E., and Yahya, H. A., *A Comparative Analysis of Design Criteria Influencing Building Material Selection Across Different Architectural Contexts*, International Journal of Sustainable Development and Planning, 18(10), 3117–3124, 2023.

Climate Positive Europe Alliance, *ESG-Verification for the EU-Taxonomy: New Construction*, 2023. Retrieved from https://is.gd/MmvQ9t on March 30, 2025.

DGNB Systems, *DGNB System - New Construction, Buildings Criteria Set. Version 2023*, 2024. Retrieved from https://www.dgnb.de/en/certification/buildings/new-construction on March 29, 2025.

Hou, K., Li, J., Liu, Y., Sun, S., Zhang, H., and Jiang, H., *KG-EGV: A Framework for Question Answering with Integrated Knowledge Graphs and Large Language Models*, Electronics, MDPI, 13(23), 4835, December 2024.

Larson, J., and Truitt, S., *GraphRAG: A New Approach for Discovery Using Complex Information*, 2024. Retrieved from https://is.gd/Q0dQpe on April 1, 2025.

Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Küttler, H., Lewis, M., Yih, W. T., Rocktäschel, T., Riedel, S., and Kiela, D., *Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks*, Advances in Neural Information Processing Systems, 33, 9459-9474, 2020.

Schneider-Marin, P., Stocker, T., Abele, O., Margesin, M., Staudt, J., Abualdenien, J., and Lang, W., *EarlyData Knowledge Base for Material Decisions in Building Design*, Advanced Engineering Informatics, Elsevier, 54, 101769, October 2022.

Wei, J., Wang, X., Schuurmans, D., Bosma, M., Ichter, B., Xia, F., Chi, E., Le, Q., and Zhou, D., *Chain-of-Thought Prompting Elicits Reasoning in Large Language Models*, Advances in Neural Information Processing Systems, 35, 24824-24837, 2022.

Yao, S., Zhao, J., Yu, D., Du, N., Shafran, I., Narasimhan, K., and Cao, Y. *ReAct: Synergizing Reasoning and Acting in Language Models*, In International Conference on Learning Representations (ICLR), 2023.

Zhang, Q., Chen, S., Bei, Y., Yuan, Z., Zhou, H., Hong, Z., Dong, J., Chen, H., Chang, Y., and Huang, X., *A Survey of Graph Retrieval-Augmented Generation for Customized Large Language Models*, arXiv preprint, 2025.