2025 European Conference on Computing in Construction

CIB W78 Conference on IT in Construction

Porto, Portugal
July 14-17, 2025

DIGITAL DECISION SUPPORT IN CONSTRUCTION: ANALYSIS OF INFORMATION REQUIREMENTS AND DATA PROVISION FOR AI-BASED SELECTION OF SUSTAINABLE BUILDING PRODUCTS

Dominik Stellmacher[1], Karina Große Lögten[2], Sebastian Theißen[1,2], Agnes Kelm[1], and Univ.-Prof. Dr.-Ing. habil. Anica Meins-Becker[1]

[1] University of Wuppertal, Germany

[2] List Eco GmbH

Abstract

The construction industry is increasingly adopting sustainable, resource-efficient practices requiring data-driven decisions. However, manual data preparation and lack of standardisation impede this transition. Certification systems like DGNB, quality seals like QNG and classification systems like the EU Taxonomy incorporate measurable sustainability criteria, highlighting the need for reliable data. This study utilises LLMs to extract data from product data sheets and Environmental Product Declarations, converting it to standardised JSON schema. By integrating Python and knowledge graphs, the structured data can be matched against sustainability criteria, serving as foundation for planners, contractors, and clients to make AI-based decisions for sustainable building products.

Introduction

The construction industry is responsible for a significant share of global energy and resource consumption, as well as greenhouse gas emissions (United Nations 2021). Against the backdrop of climate change and international sustainability goals such as the UN Sustainable Development Goals (UN-SDGs) (United Nations 2015), a consistent orientation towards environmentally friendly, resource-efficient, and healthy construction practices is indispensable. Building products play a central role in this context, as their environmental impacts over the entire lifecycle of a building substantially contribute to its sustainability performance (Sandanayake 2022).

Contribution of Materials and Building Products to the Sustainable Development Goals (SDGs)

In view of the increasing urgency to combat climate change and utilise resources more efficiently, the United Nations adopted the UN-SDGs in 2015 (United Nations 2015). These 17 goals for sustainable development form a global action plan to end poverty, protect the planet, and ensure prosperity for all. In the context of construction, materials and building products are of central importance for achieving several of these goals, particularly SDG 13 ("Climate Action"), which demands a rapid reduction in greenhouse gas emissions. The selection and use of sustainable materials and building products can significantly reduce energy demand and lower CO_2 emissions over the entire lifecycle of a building (Sandanayake 2022). Moreover, the use of regionally available materials contributes to a reduction in transport-related emissions (Omer and Noguchi 2020).

Figure 1: The Impact of Materials and Building Products on the SDGs (World Green Building Council)

Figure 1 illustrates the impact of materials and construction products on the Sustainable Development Goals (SDGs). Among the most relevant are (Omer and Noguchi 2020):

- SDG 3 (Good Health and Well-being):
  Low-emission materials reduce pollutants (e.g. VOCs). Certifications such as Cradle to Cradle Certified® or Blue Angel ensure a healthy indoor environment.
- SDG 7 (Affordable and Clean Energy):
  Energy-efficient materials (e.g. high-performance insulation) lower heating and cooling demands and promote a reliable, sustainable energy supply.
- SDG 8 (Decent Work and Economic Growth) and SDG 10 (Reduced Inequalities):
  Labels such as the Concrete Sustainability Council (CSC) or Forest Stewardship Council (FSC 2025) enhance transparency along the supply chain and consider social and ecological criteria.
- SDG 12 (Responsible Consumption and Production):
  Recycled or reusable construction materials contribute to the responsible use of natural resources.
- SDG 13 (Climate Action):
  Building products with low Global Warming Potential (GWP) reduce greenhouse gas emissions during production, use, and disposal.

These interrelations underscore the importance of selecting building products not only in terms of their technical properties and economic factors, but also from social and ecological perspectives. A well-founded understanding of life cycle emissions, resource consumption, and potential pollutant loads is indispensable.

Role of Certification Systems, Seals and European Standards

Various certification systems and standards establish a structured framework for assessing the sustainability performance of buildings, materials, and construction products. In Germany, this particularly includes the EU Taxonomy, the German Sustainable Building Council (DGNB) system, and the Quality Seal for Sustainable Building (QNG). Additionally, international certification systems such as LEED (USA) and BREEAM (United Kingdom) are also implemented in Germany.

EU Taxonomy

The EU Taxonomy (European Union 2020) defines clear criteria by which economic activities can be classified as environmentally sustainable, creating equal market conditions and targeted financial flows for climate-friendly projects (DGNB, DK-GBC, GBCe, ÖGNI). The criteria are divided into six overarching environmental objectives (DGNB System and CPEA 2022):

- Climate change mitigation
- Climate change adaptation
- Sustainable use and protection of water and marine resources
- Transition to a circular economy, waste prevention, recycling
- Pollution prevention and reduction
- Protection and restoration of biodiversity and ecosystems

By selecting resource-efficient and low-emission building materials, objectives relating to a circular economy, waste prevention, and pollution reduction are particularly addressed. In addition to selecting sustainable building materials, further measures are necessary to achieve the other objectives. For example, to meet the goal of climate adaptation, climate risks must be identified through a climate risk analysis and, if necessary, appropriate adaptation solutions must be implemented. In this way, the fulfilment of all relevant taxonomy criteria in the construction context is ensured.

DGNB

The DGNB system for new construction version 2023 integrates 29 criteria from the domains of ecology, economy, and sociocultural aspects into six overarching thematic fields (ENV, ECO, SOC, TEC, PRO, SITE) (DGNB System 2023). In the thematic field “Ecological Quality” (ENV), the three main criteria target climate protection, reduction of potentially harmful building products, and responsible resource extraction, with indicators such as Global Warming Potential (GWP), Non-renewable primary energy requirement (PENRT), Photochemical Ozone Creation Potential (POCP), Eutrophication Potential (EP), Acidification Potential of Soil and Water (AP), and Freshwater Consumption (FW) being decisive. “Economic Quality” (ECO1.1) considers all costs incurred over the building's lifecycle, thereby opening up early optimisation potentials. Under “Sociocultural and Functional Quality” (SOC), aspects such as indoor air quality, comfort, and user satisfaction are evaluated, while the area of “Technical Quality” (TEC) focuses on circular construction (TEC1.6). “Process Quality” (PRO) deals with planning, execution, and documentation of the construction process, and “Site Quality” (SITE) examines infrastructure, environmental conditions, and connectivity. For the Life Cycle Assessment (LCA) according to the calculation rules of the DGNB, which refer directly to the QNG regulations, the life cycle phases A1–A3 (production), B4 (replacement) and B6 (operational energy), as well as C3–C4 (disposal) are primarily utilised. Already in the early planning stages, a targeted selection of building materials decisively influences the fulfilment of central criteria – for example, the reduction of emissions and pollutant loads – thereby simultaneously promoting efficient and sustainable construction practices.

QNG

The QNG was initiated within the framework of the German Federal Funding for Efficient Buildings (BEG) and sets limits for GWP and PENRT. Furthermore, for non-residential buildings, it includes requirements regarding sustainable material sourcing, pollutant avoidance, accessibility, natural hazards at the site, and green roofs. This quality seal enables a transparent evaluation of ecological and sociocultural building qualities and thus promotes a consistently sustainable planning concept (German Housing, Urban Development and Building).

Challenges in Implementation

Although certification systems such as DGNB and QNG, together with the EU Taxonomy, provide a clear framework for sustainable decision-making processes, practical integration into the planning and construction process is complex. Data on building products often exist in different, sometimes unstructured formats and sources, for instance in Product Data Sheets (PDS), Environmental Product Declarations (EPDs) according to DIN ISO 14025, Safety Data Sheets (SaDS), Sustainable Data Sheets (SuDS), or manufacturer catalogues. The manual collection, verification, and updating of this information is time-consuming and prone to error; furthermore, disparate data structures can lead to information gaps (Hofstadler and Motzko 2021). Current DGNB study results show that a lack of a data basis is a central problem for market participants. Only one of 62 examined projects exhibited full taxonomy compliance, partly due to insufficient data for the required proof. These deficiencies are particularly evident in the criteria “Climate Change Mitigation” and “Climate Change Adaptation”. This results in a significant need for catch-up in information processing, especially concerning existing but unstructured data throughout the entire lifecycle of buildings (DGNB 2021). A central digital recording and automated processing of product information is therefore indispensable to create reliable and comparable foundations for certification and standard compliance. In this way, a consistent data basis for assessing the sustainability performance of materials and building products can be established (buildingSMART Switzerland and BAUEN DIGITAL SCHWEIZ 2024), which meets the stringent requirements of national systems (e.g., DGNB, QNG) as well as international regulations such as the EU Taxonomy.

Objective of the Study

The research aims to develop an AI-supported decision support system for the sustainable selection of building products. To this end, the information requirements of various certification systems will first be transformed into machine-readable formats, and building products will be systematically classified according to the Standardised Service Specifications for Construction (STLB). Since the STLB has provided standardised descriptions of construction services since the 1960s, it enables a uniform classification and detailed recording of building products (STLB-Bau 2025). Modern LLMs automatically extract relevant data from unstructured sources and structure it accordingly. Subsequently, the obtained information is semantically transferred into an ontology-based knowledge graph (KG), enabling efficient data management and rapid queries. Through an integrated validation loop, the extracted values and units are repeatedly verified – with a confidence score provided – to enhance data accuracy. Using a developed GraphRAG pipeline (Graph-based Retrieval Augmented Generation), in which an LLM accesses the information stored in the Knowledge Graph (KG), architects, construction professionals, sustainability auditors, and building owners receive precise, data-driven responses that support them in the reliable selection of sustainable building materials.

Analysis and Preparation of Information Requirements for the Selection of Sustainable Building Products

The foundation for AI-supported, semantically grounded information acquisition and processing lies in the systematic identification and structuring of relevant criteria. The aim is to prepare requirements derived from DGNB, QNG, and the EU Taxonomy such that they can be automatically verified.

Methodology for Determining Information Requirements

Within the NaConBau research project (University of Wuppertal - Digital planning, construction and operation 2024), a methodology was developed to systematically derive relevant information requirements for sustainable building products. The starting point was the criteria catalogues of DGNB (DGNB System 2023) and QNG (QNG Sustainable Building Certification 2024), as well as the stipulations of the EU Taxonomy.

Procedure:

1. Categorisation by Trades, Product Groups, and Material Types:
   Based on the STLB, building products are assigned to various trades, such as masonry or concrete work, then clustered into product groups like masonry units or masonry mortar, which in turn are differentiated into product types such as sand-lime brick or masonry stone.

Table 1: Clustering of Information into four Principal Categories (Alam Bhuiyan and Hammad 2023)

Category	Description	Typical Information	Possible Sources
Technical Information	Mechanical, physical, and building physics properties, essential for construction requirements	Dimensions (LxWxH), weight, bulk density (class), compressive strength, thermal conductivity, U-value, sound insulation, fire resistance class, moisture behaviour, consumption, yield, processing temperature, service life, product description, application area	PDS, performance declaration, EPD
Environmental Information	Environmental impacts over the entire lifecycle based on standardised evaluation methods (e.g., according to DIN EN 15804+A2)	Declaration number, declared unit, validity, standardisation basis (EN 15804+A1/A2), environmental indicators (e.g., GWP, AP, EP, POCP), resource use (e.g., PENRT), recyclability, primary & secondary resource use	EPD, IBU database, Ökobaudat, SuDS
Economic Information	Full lifecycle costs for holistic economic evaluation	Material costs, transport costs, maintenance costs, disposal costs	Cost calculations (internal), manufacturer information, calculation databases
Social Information	Aspects related to working conditions, health, safety, and certification	Pollutant content (e.g., VOC content), certifications (e.g., Blue Angel, FSC, WiN Fair Stone, EMICODE)	SuDS, WECOBIS, Sentinel Haus Institute

Table 2: Excerpt of Information Requirements Exemplified by Masonry Work, (DGNB System 2023; QNG Sustainable Building Certification 2024; DGNB, DK-GBC, GBCe, ÖGNI; DGNB System and CPEA 2022)

	Eco-Balance[1]		Pollutants						Docu-ment, Sources
Product Type	DGNB V23 (ENV 1.1)	QNG	DGNB V23 (ENV1.2)				QNG	EU Taxonomy	
			QL1	QL2	QL3	QL4			
Masonry (e.g. sand-lime brick)	GWP, PENRT, AP, EP, POCP	GWP, PENRT	None	None	None	None	None	None	EPD, PDS
Disper-sion Insulation Adhesive	None	None	VOC < 40 g/l	VOC < 40 g/l	VOC < 40 g/l	• VOC < 40 g/l • Biocides: Justification for the use of film-preserved products, incorporat-ing UBA fact sheets	• VOC ≤ 30.0 g/l (water-based formulation) • Declaration of biocidal active substances • Lead compounds ≤ 0.10 %	• SVHC substances ≤ 0.10 % • CMR substances category 1A or 1B • ≤ 0.10 %	SaDS, SuDS, PDS

2. Identification of Relevant Criteria:
   For each product type, in addition to recording the product and manufacturer names, data from the four main categories of sustainable product selection – technical, economic, ecological, and sociocultural (Alam Bhuiyan and Hammad 2023) – are captured to ensure unambiguous identification and description.
   In this context, relevant product information was identified as shown in Table 1, divided into the four main categories to comprehensively capture the diversity of sustainability assessment.
3. Assignment of Document Sources:
   The requirements were linked to typical document sources such as Product Data Sheets, Environmental Product Declarations, Safety Data Sheets, and Sustainability Data Sheets (cf. Table 1). Additionally, external databases, for example the DGNB Navigator, WECOBIS, or the Sentinel Haus Institute, can be utilised.
4. Tabular Preparation:

All identified criteria were recorded in tables (cf. Table 2) to clearly define which information is needed for which material group and from which sources.

JSON Representation of Requirements and Extracted Product Information

To render both the sustainability requirements and the product information extracted from unstructured sources machine-readable, they are converted into a structured JSON format (Friesen 2019). This format facilitates the alignment of product information with the sustainability requirements. Figure 2, for example, illustrates the sustainability requirements of the DGNB quality levels (QL) for a dispersion insulation adhesive.

This enables a clear hierarchical structuring of the data, flexible supplementation with new criteria, as well as seamless integration into knowledge graphs or other systems for further processing.

This facilitates the transfer into KGs as well as the use of LLMs for context-sensitive processing.

[1] Under the EU Taxonomy framework, no specific threshold values (e.g., for GWP and PENRT) are mandated to be met within a life cycle assessment (LCA). However, for new constructions exceeding 5,000 m² gross floor area (GFA), the completion of an LCA is mandatory, with particular emphasis on the disclosure of GWP values. (DGNB System and CPEA 2022).

```
"certificationSystems": [
  "DGNB",
  "23",
  "criteria": [
   {
    "name": "ENV1.2",
    "constructionProductTypes": [
     {
      "name": "Dispersion Insulation Adhesive",
      "qualityLevels": [
       {"level": "QL1", "requirements": [ {"type": "VOC_Content", "name": "voc_content", "operator": "<", "value": 40, "unit": "g/L", "datatype": "float"}]},
       {"level": "QL2", "requirements": [ {"type": "VOC_Content", "name": "voc_content", "operator": "<", "value": 40, "unit": "g/L", "datatype": "float" }]},
       {"level": "QL3", "requirements": [ {"type": "VOC_Content", "name": "voc_content", "operator": "<", "value": 40, "unit": "g/L", "datatype": "float" }]},
       {"level": "QL4", "requirements": [{"type": "VOC_Content", "name": "voc_content", "operator": "<", "value": 40, "unit": "g/L", "datatype": "float"},
          {"type": "Documentation", "name": "film_protected_products_justification", "description": "Justification of film-protected products, incorporating UBA fact sheets"}
         ]}
      ] }
    ]}
 ]},
```

Figure 2: JSON Excerpt of Requirements for Building Products: A Case Study of DGNB Quality Levels for Dispersion Insulation Adhesive

Information Acquisition and Processing of Building Product Information

In addition to defining the information requirements, the efficient collection of data forms the basis for an AI-supported decision support system for selecting sustainable building products. The optimized process of data collection and processing is depicted in Figure 3 and explained below, with a particular focus on automation, validation, and transparency.

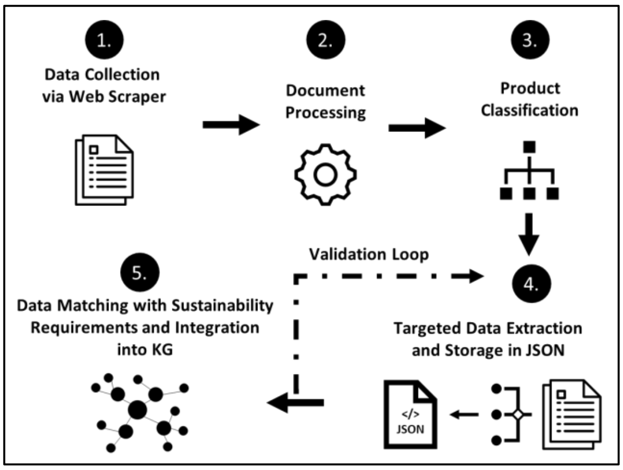

Figure 3: Process Diagram from Data Acquisition to Integration into the Knowledge Graph

1. Web Scraping: Automated Data Collection

Using Python scripts and libraries such as *Selenium* and *BeautifulSoup*, relevant documents are automatically collected directly from manufacturer websites. The locally stored PDF documents are classified according to document types based on URL or HTML structures and prefixed with specific tags such as “PDS_” or “EPD_”, which form the basis for subsequent data extraction. This approach ensures targeted information extraction and reduces the number of tokens required for LLM queries.

2. Document Processing: Extraction of Text and Tables

The collected PDF documents are pre-processed using Python and libraries such as pdfplumber and camelot. While pdfplumber extracts continuous text from the PDFs, camelot converts tables into formats such as .xlsx or .csv. Notably, important environmental indicators from the EPD documents can be directly obtained from the CSV files without invoking the LLM for the extraction process, thereby conserving resources and minimizing potential error sources.

3. Product Classification Using LLM

The precise classification of building products is essential for the efficient and targeted extraction of product-specific information. In this step, OpenAI’s GPT-4o-mini model (OpenAI 2025) is employed as an LLM via API, integrated into the existing Python application using the LangChain library. LangChain is an open-source Python framework that significantly facilitates the development of applications with LLMs (Langchain 2025).

CoT-Prompt:

```
product_type_descriptions =
    { "Lime-sand brick": "A masonry unit made from a mixture of lime and sand, which after autoclaving
    typically has precise dimensions (or smooth face) and offers high compressive strength while providing
    good thermal insulation.",
    "Brick": "Traditionally fired masonry unit made from clay or loam, which exhibits high compressive
    strength, resistance to moisture, and proven suitability for both load-bearing and non-load-bearing
    walls.",
    "Standard mortar": "A standard masonry mortar made from cement, lime, and sand, developed for
    general masonry work and characterized by good adhesion and workability, as well as sufficient
    strength and flexibility.",
    "Dispersion insulation adhesive": "A special adhesive based on a dispersion, optimized for bonding
    insulation materials to various substrates. It is distinguished by high adhesion, flexibility, and long-
    lasting resistance to moisture.",
    "Gypsum plasterboard": "A drywall panel consisting of a gypsum core sheathed on both sides with
    paper. It enables fast and clean interior finishing work, offers sound insulation and fire resistance, and is
    easy to work with."
}

# System prompt for product type determination with visible chain-of-thought
system_prompt_type = (
    "You are an expert assistant for building materials. "
    "You will receive a document text and a suggestion for a product type.\n"
    "Think step by step and explain in detail:\n"
    "1. Analyze the document text.\n"
    "2. Compare the description of the suggested product type with the product types from the list:\n"
    f"{product_type_descriptions}\n"
    "3. Justify whether the suggested type fits or not.\n"
    "4. If another type fits better, explain why.\n"
    "First, provide the justification before clearly naming the final product type at the end."
)
```

Figure 4: Chain-of-Thought Prompt for the Classification of Building Product Types

A central element of this classification process is the application of the Chain-of-Thought (CoT) method, as depicted in Figure 4. CoT is a technique in the field of LLMs where the model is guided to elaborate its thought processes and intermediate steps in detail. By presenting the reasoning step by step, CoT enables transparent and comprehensible decision-making. Furthermore, this method helps to enhance performance in complex tasks that require multi-step reasoning and logical inference (Wei et al. 2022).

4. Targeted Data Extraction per Product Type

Specialized Python scripts are employed during data extraction to retrieve relevant product information – such as density, consumption values, or emission data – from unstructured sources using targeted prompts. Initially, the document content is passed to a large language model (LLM) that, based on a system-defined prompt, extracts the desired data (e.g., value and unit) and returns it in JSON format. To verify the accuracy of the extracted data, a validation loop is implemented. This loop formulates a validation prompt in which the LLM is asked to check the original text to confirm whether the extracted value and unit are correct. If the LLM simply returns "OK," the information is deemed correct; if a response in the format "value|unit" is provided, it is interpreted as a correction. The validation iteratively repeats this process until the same result is obtained twice in succession. Once confirmed, the output is assigned a confidence score of 1.0, signaling to the user that the information is considered absolutely correct. Should no matching validation occur after the maximum number of iterations (10), the most recent result is accepted with a confidence score of 0.5. This score reflects a degree of uncertainty while substantially reducing the likelihood of error, as random misextractions are statistically unlikely to occur consecutively. Thus, the system not only provides the desired data but also offers a reliable estimate of data accuracy, forming the basis for well-founded decisions.

5. Data Matching and Integration into Knowledge Graph

Finally, the processed product data is integrated into the knowledge graph (Neo4j). A Python script automatically verifies whether the building products meet the defined quality levels (QL) according to DGNB, as well as the requirements of QNG and the EU taxonomy – for example, validating that the VOC content is ≤ 40 g/l. Data matching is performed using the JSON files of the building products and the sustainability requirements, as exemplarily shown in Figure 4. Upon successful validation, a relationship ("FULFILS_REQUIREMENTS_OF") is established between the product node and the corresponding quality level. Figure 5 displays a segment of the knowledge graph, illustrating how building products are linked with their color-coded properties, fulfilled sustainability requirements, and DGNB quality levels. In parallel, the extracted information is categorized according to Table 1 into "Technical" (gray), "Ecological" (light green), "Economic" (dark blue), and "Sociocultural" (purple). Additionally, the filenames of the original documents from which the information was extracted are stored within the document nodes, thereby facilitating easy access for manual reviews and enhancing transparency. Structured storage in an ontology-supported knowledge graph enables complex queries and ensures the overall transparency of the process.

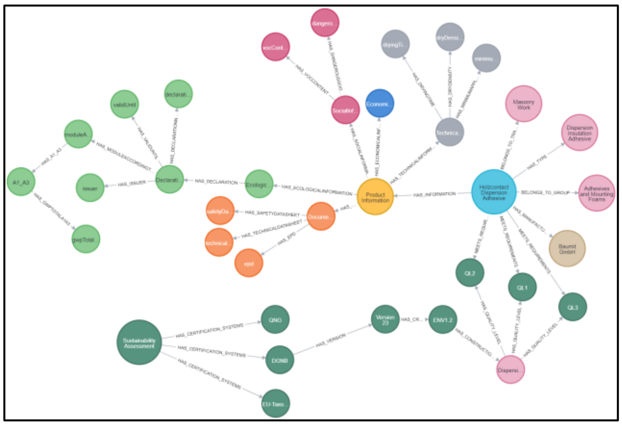

Figure 5: Excerpt from the Knowledge Graph Illustrating the Link between Requirements and Building Products

Results

To validate the system, eight building products – including two dispersion insulation adhesives, a sand-lime brick, a cellular concrete block, a multi-mortar, a gypsum board, a roof rafter insulation, and a concrete product – were used to demonstrate the entire AI-supported workflow. Unstructured data from technical product datasheets were automatically extracted, transformed into a standardized JSON format, and integrated into a knowledge graph (Neo4j). A chain-of-thought prompt ensured the correct assignment of products to their respective product types. For products with multiple dimensional variants, a total of 29 JSON objects were created, encompassing 208 technical and social data fields. The validation of these fields was conducted using confidence scores, achieving an average score of 0.986 – with only three fields exhibiting uncertainty that required manual review. Ecological data were extracted directly from EPD documents, while economic information could not be determined due to missing price data.

The automated matching of data is performed based on the JSON files of the building products and the sustainability requirements – as shown in Figure 4 – and, upon successful validation, results in the creation of a relationship ("FULFILS_REQUIREMENTS_OF") between the product node and the corresponding quality level. Additionally, the implemented GraphRAG pipeline enables an LLM to access the structured information stored in the knowledge graph, translating human-formulated queries into machine-readable Cypher queries, retrieving relevant information, and presenting it in clear language. This data-driven, AI-supported decision-making process assists architects, construction professionals, sustainability auditors, and building owners in making well-founded and transparent decisions regarding the selection and substantiation of sustainable building products.

Sample queries and their results are as follows:

- **Example Query 1:**

  Question: Which sand-lime brick *(product type)* with a compressive strength > 12 N/mm² and a width of 240 mm *(technical)* has the lowest total GWP value for phase A1-A3 *(ecological)*?

  Generated Cypher Query:

  MATCH (bp:BuildingProduct)-
  [:BELONGS_TO_BUILDING_PRODUCT_TYPE] →
  (bpt:BuildingProductType {label: "Sand Lime Brick"})
  MATCH (bp)- [*] → (df:CompressiveStrength)
  MATCH (bp)- [*] → (gwp:GWP_Total_A1_A3)
  WHERE toFloat(df.value) > 12 AND df.unit = "N/mm²"
  RETURN bp.label AS BuildingProduct, df.value AS CompressiveStrength, df.unit AS Unit, gwp.value AS GWP_A1_A3_total_Value, gwp.unit AS GWP_A1_A3_total_Unit
  ORDER BY toFloat(gwp.value) ASC
  LIMIT 1

  Answer: *The sand-lime brick that meets all the criteria is "Product A." It has a compressive strength of 13.8 N/mm² and a total GWP value of 0.488 kg CO_2-equivalent.*

- **Example Query 2:**

  Question: Which dispersion insulation adhesive *(product type)* meets DGNB Quality Level 3 *(social – due to its VOC content)* and has the lowest material costs per unit (economic)? (Note: A separate material price list is available.)

  Generated Cypher Query:

  MATCH (bp:BuildingProduct)-
  [:BELONGS_TO_BUILDING_PRODUCT_TYPE] →
  (bpt:BuildingProductType {label: "Dispersion Insulation Adhesive"})
  MATCH (bp)-[r:FULFILS_REQUIREMENTS_OF] →
  (qs:QualityLevel {label: "QS3"})
  RETURN bp.label AS BuildingProduct, type(r) AS RelationshipType, qs.label AS QualityLevel

  Answer: The pipeline determined that the "4th Dispersion Insulation Adhesive PCI" – based on the available cost data – fulfills DGNB Quality Level QS3 and, with material costs of 8 €/m² (compared to 12 €/m² for another candidate), exhibits the lowest cost per unit.

In summary, the evaluation demonstrates that the presented workflow can automatically and accurately analyze extensive product data. The combination of LLM-supported data extraction, validation-based error mitigation, and integration into a knowledge graph enables data-driven, AI-supported decision-making. The GraphRAG pipeline translates complex, human-formulated queries into Cypher queries, retrieves relevant information, and presents it in clear language – a method that supports complex multi-hop queries across technical, ecological, and sociocultural criteria. This approach optimally assists architects, construction professionals, sustainability auditors, and building owners in making well-founded decisions regarding the selection and substantiation of sustainable building products.

Conclusion

Validation of the extracted data reveals that the entire process – based on an average confidence score of 0.986 – provides a highly reliable foundation for structuring unstructured information via the developed pipeline. This high score demonstrates that automated extraction and classification are, in most cases, precise and require only minimal manual review. Thus, the workflow forms a solid basis for data-driven decisions in construction. The integrated GraphRAG pipeline enables an LLM to access, in real time, the information stored in the knowledge graph (Neo4j) and translate human queries into machine-readable Cypher queries. Through multi-hop queries (Shavaki et al. 2024), complex interrelations among technical, ecological, and sociocultural criteria can be identified and evaluated. Additionally, where available, economic factors – such as proprietary material lists or calculation approaches – can be incorporated, facilitating a comprehensive assessment and comparison of building products.

It should be noted that, similar to other AI-based systems such as ChatGPT, the automated results here are not error-free. The file paths to the original documents stored within the knowledge graph provide an essential means to manually review and, if necessary, correct the automated outputs. This critical review ensures that users can consistently make well-founded and verifiable decisions.

Overall, the presented workflow supports architects, construction professionals, sustainability auditors, and building owners by providing a transparent and data-driven basis for decision-making. It complements expert technical knowledge by enabling rapid preselection and evaluation of relevant product information, thereby significantly facilitating the selection and substantiation of sustainable building products.

Acknowledgments

This research was carried out within the framework of the NaConBau project at the University of Wuppertal in cooperation with List Eco GmbH and was funded by the Federal Ministry of Education and Research (BMBF).

References

Alam Bhuiyan, Mohammad Masfiqul/Hammad, Ahmed (2023). A Hybrid Multi-Criteria Decision Support System for Selecting the Most Sustainable Structural Material for a Multistory Building Construction. Sustainability 15 (4), 3128. https://doi.org/10.3390/su15043128.

buildingSMART Switzerland/BAUEN DIGITAL SCHWEIZ (2024). Digitally networked building product data as the basis for circularity. Requirements and solutions, 35. (accessed 1/7/2024).

DGNB (2021). Study on EU taxonomy: companies only partially prepared for requirements. Available online at https://www.dgnb.de/en/making-the-most-of-dgnb/newsroom/press/article/study-on-eu-taxonomy-companies-only-partially-prepared-for-requirements (accessed 1/7/2025).

DGNB System (2023). NEW CONSTRUCTION, BUILDINGS CRITERIA SET. Version 2023.

Available online at criteria catalog new building (accessed 07.0.2025).

DGNB System/CPEA (2022). ESG VERIFICATION FOR THE EU TAXONOMY: NEW CONSTRUCTION - CLIMATE CHANGE MITIGATION. Available online at https://www.dgnb.de/filestorages/Downloads_unprotected/dokumente/esg-verification-climate-mitigation-new-construction-requirements-en.pdf (accessed 1/7/2025).

DGNB, DK-GBC, GBCe, ÖGNI. EU TAXONOMY STUDY. Evaluating the market-readiness of the EU taxonomy criteria for buildings. Available online at https://gbce.es/wp-content/uploads/2023/08/gbcs_eu_taxonomy_market_readyness_study.pdf (accessed 1/7/2025).

European Union (2020). REGULATION (EU) 2020/852 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL. Available online at https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32020R0852 (accessed 1/7/2025).

Friesen, Jeff (2019). Java XML and JSON. Berkeley, CA, Apress.

German Housing, Urban Development and Building. Quality seal for sustainable buildings (QNG). New construction and modernization of residential and non-residential buildings 2023, 32. Available online at https://www.bmwsb.bund.de/SharedDocs/downloads/Webs/BMWSB/DE/publikationen/bauen/qng-neubau-und-modernisierung-von-wohn-und-nichtwohngebaeuden.pdf?__blob=publicationFile&v=2 (accessed 1/7/2025).

Hofstadler, Christian/Motzko, Christoph (Eds.) (2021). Agile Digitalisierung im Baubetrieb. Grundlagen, Innovationen, Disruptionen und Best Practices. Wiesbaden/Heidelberg, Springer Vieweg.

Langchain (2025). Introduction. Available online at https://python.langchain.com/docs/introduction/ (accessed 1/7/2025).

Omer, Mohamed A.B./Noguchi, Takafumi (2020). A conceptual framework for understanding the contribution of building materials in the achievement of Sustainable Development Goals (SDGs). Sustainable Cities and Society 52, 101869. https://doi.org/10.1016/j.scs.2019.101869.

OpenAI (2025). GPT-4o-mini. Available online at https://platform.openai.com/docs/models#gpt-4o-mini (accessed 1/7/2025).

QNG Sustainable Building Certification (2024). Appendix 2 - QNG CRITERIA LIST. Available online at https://www.qng.info/app/uploads/2024/11/QNG_Handbuch_Anlage-2-Systemanforderungen_v1-3_en.pdf (accessed 1/7/2025).

Sandanayake, Malindu Sasanka (2022). Environmental Impacts of Construction in Building Industry—A Review of Knowledge Advances, Gaps and Future Directions. Knowledge 2 (1), 139–156. https://doi.org/10.3390/knowledge2010008.

Shavaki, Mahdi Amiri/Omrani, Pouria/Toosi, Ramin/Akhaee, Mohammad Ali (2024). Knowledge Graph Based Retrieval-Augmented Generation for Multi-Hop Question Answering Enhancement. In: 2024 15th International Conference on Information and Knowledge Technology (IKT), Isfahan, Iran, Islamic Republic of, 24.12.2024 - 26.12.2024. IEEE, 78–84.

STLB-Bau (2025). Service areas - STLB-Bau tender texts. Available online at https://www.stlb-bau-online.de/Ausschreibungstexte/Leistungsbereiche/1 (accessed 1/7/2025).

United Nations (2015). THE 2030 AGENDA FOR SUSTAINABLE DEVELOPMENT. Available online at https://sustainabledevelopment.un.org/content/documents/21252030%20Agenda%20for%20Sustainable%20Development%20web.pdf (accessed 1/7/2025).

United Nations (2021). 2021 GLOBAL STATUS REPORT FOR BUILDINGS AND CONSTRUCTION. Towards a zero-emissions, efficient and resilient buildings and construction sector. Available online at https://globalabc.org/sites/default/files/2021-10/GABC_Buildings-GSR-2021_BOOK.pdf (accessed 1/7/2025).

University of Wuppertal - Digital planning, construction and operation (2024). NaConBau. Sustainable controlling tool for construction work. Available online at https://dpbb.uni-wuppertal.de/en/research/current-research-projects/naconbau/ (accessed 1/7/2025).

Wei, Jason/Wang, Xuezhi/Schuurmans, Dale/Bosma, Maarten/Ichter, Brian/Xia, Fei/Chi, Ed/Le Quoc/Zhou, Denny (2022). Chain-of-Thought Prompting Elicits Reasoning in Large Language Models. Available online at http://arxiv.org/pdf/2201.11903.

World Green Building Council. Thought Leadership. Green building: Improving the lives of billions by helping to achieve the UN Sustainable Development Goals. Available online at https://worldgbc.org/article/green-building-improving-the-lives-of-billions-by-helping-to-achieve-the-un-sustainable-development-goals/ (accessed 1/7/2025).