DOI: https://doi.org/10.30525/2256-0742/2026-12-5-78-85

AN LLM-BASED FRAMEWORK FOR AUTOMATED RECOGNITION AND STRUCTURING OF CONSTRUCTION COST ESTIMATES

Oskars Bagants[1], Raita Rollande[2], Juris Klonovs[3]

Abstract. The construction industry continues to encounter considerable inefficiencies associated with the manual interpretation and processing of heterogeneous cost estimation documentation, which is characterised by a paucity of background data. The present study puts forward a Large Language Model (LLM)-based framework for the automated recognition, semantic structuring, and enrichment of construction cost estimates originating from unstandardised Excel spreadsheets. The proposed framework uses a combination of domain-specific data annotation, machine learning-assisted document understanding and semantic similarity retrieval to convert semi-structured procurement and estimation documents into standardised, machine-readable datasets. A large corpus of historical construction estimates was collected, organised and manually labelled to create a training dataset for the LLM-based recognition pipeline. An API-oriented architecture automatically parses, classifies and converts uploaded spreadsheets into structured entities, including estimate positions, quantities, units, unit costs and total costs. The framework also incorporates a semantic similarity matching mechanism that compares newly uploaded estimates with historical priced estimates stored in the reference database. Consequently, the system can automatically suggest historically relevant unit prices and related CO_2 indicators for similar construction positions. Experimental validation has demonstrated that the proposed framework achieves a semantic position-matching accuracy of over 85%, enabling the automatic completion of 50–70% of newly uploaded estimates. Furthermore, the solution reduces the effort required for manual estimation by around 70% compared to traditional workflows. The findings suggest that using LLM-assisted recognition and historical knowledge retrieval can substantially improve the productivity, consistency and sustainability of construction cost management and procurement processes.

Keywords: construction cost estimation, large language models, semantic similarity matching, structured data extraction, sustainability, natural language processing.

JEL Classification: C55, C88, D83, L74, O33, Q01

1. Introduction

The construction industry still relies heavily on manually prepared, diverse cost estimation documents, particularly during the procurement, bidding and tendering processes (Abanda et al., 2017; Deza et al., 2022). Although automation of quantity take-off processes has advanced significantly, most commercial cost estimation software still relies on manual intervention by estimators to link building elements to the relevant work items and/or provide essential information missing from design models that is necessary for accurate cost estimation. This missing data is usually extracted from construction specifications that complement the design documentation. Automating the extraction and interpretation of information from construction specifications could substantially reduce the time and financial resources required for estimation activities. It would also minimise the need for manual input in cost calculation procedures and lower the likelihood of human-related estimation errors (Akanbi and Zhang, 2020; Akanbi and Zhang, 2021).

Received on: July 29, 2026 **Accepted on:** October 12, 2026 **Published on:** November 02, 2026

[1] BIKUS SIA, Latvia *(corresponding author's)*
E-mail: oskars@bikus.lv
ORCID: https://orcid.org/0009-0002-2575-9614
[2] Ventspils University of Applied Sciences, Latvia
ORCID: https://orcid.org/0000-0003-3629-2677
[3] Ventspils University of Applied Sciences, Latvia;
BirgerMind SIA, Latvia
ORCID: https://orcid.org/0000-0002-4615-3163

This is an Open Access article, distributed under the terms of the Creative Commons Attribution CC BY 4.0

Construction cost estimates are often shared in spreadsheet format, but these spreadsheets have inconsistent structures, different terminology and non-standardised position descriptions. This results in substantial manual effort being required to interpret, structure, compare and complete such documents, leading to inefficiencies, increased operational costs and a higher risk of human error.

Recent advances in Artificial Intelligence (AI), Natural Language Processing (NLP), machine learning and Large Language Models (LLMs) have greatly enhanced the potential for automated information extraction, semantic interpretation and knowledge representation in the Architecture, Engineering and Construction (AEC) sector. Existing research has demonstrated the applicability of AI-driven approaches in areas such as BIM-assisted construction management, automated specification and contract analysis, the semantic classification of Bills of Quantities (BoQs), cost estimation based on ontologies, and intelligent construction information retrieval systems (Abanda et al., 2017; Deza et al., 2022; Akanbi and Zhang, 2020; Akanbi and Zhang, 2021; Shen and Issa, 2010; Jadid and Idrees, 2007). In particular, NLP and machine learning techniques have demonstrated significant promise in interpreting unstructured or semi-structured documentation of construction projects, thereby enhancing interoperability between disparate data sources (Deza et al., 2022; Akanbi and Zhang, 2020). Nevertheless, despite the substantial progress made in BIM-integrated estimation systems and the processing of semantic construction data, relatively little research has addressed the automated recognition, restructuring and semantic normalisation of construction cost estimates originating from heterogeneous, spreadsheet-based procurement and tendering documentation. Existing approaches are often limited by their reliance on predefined BIM structures, rigid classification systems or narrowly defined document formats. This restricts their usefulness for real-world enterprise estimation workflows, which are characterised by highly variable Excel-based cost estimation datasets (Abanda et al., 2017; Shen and Issa, 2010; Jadid and Idrees, 2007).

The present study addresses this research gap by proposing an LLM-based framework capable of recognising, structuring, and semantically enriching construction cost estimates uploaded in arbitrary Excel formats. The novelty of the proposed framework lies in the integration of three interconnected processes:

a) Automated spreadsheet recognition and structuring;

b) semantic similarity matching using historical estimate datasets;

c) sustainability-oriented enrichment through CO_2 indicator association.

The proposed framework automatically transforms unstructured spreadsheet data into standardised, machine-readable entities, such as JSON and XML data feeds. These are suitable for enterprise-level cost management systems. Furthermore, by comparing newly uploaded estimate positions with historical databases, the framework can provide semi-automated pricing recommendations and sustainability analyses in the early stages of procurement and bidding processes.

The primary objective of this research is to evaluate the potential of using LLM-based semantic recognition methods to improve construction cost estimation workflows and reduce manual processing effort. The study also examines the potential of standardising datasets to improve the consistency, operational productivity and sustainability-oriented decision support of historical estimate retrieval mechanisms.

2. Related Work

The automation of construction cost estimation has been a popular research topic for over two decades, particularly in the context of Building Information Modelling (BIM), automated quantity take-off, ontology-based data structuring and semantic information extraction (Sun and Park, 2020; Elmousalami et al., 2025). Nevertheless, despite substantial progress in BIM-assisted estimating and automated quantity extraction, the construction industry still relies heavily on manually prepared, heterogeneous bills of quantities (BoQs), procurement spreadsheets and specification documents. Current workflows often necessitate the extensive manual interpretation, restructuring and semantic alignment of construction cost information prior to reliable cost estimation and procurement analysis.

Initial research focused primarily on automating quantity take-off and structural quantity extraction directly from CAD and BIM environments. Jadid and Idrees (2007) proposed an interactive automation algorithm for generating bills of quantities from AutoCAD drawings, employing Industry Foundation Classes (IFC) and DXF-based geometry recognition. Their study demonstrated that the automated extraction of structural quantities could improve accuracy and reduce the effort required for manual estimation. However, the proposed approach relied heavily on geometric consistency and structured CAD environments, which limited its applicability to the heterogeneous spreadsheet-based estimation documents commonly encountered in procurement workflows.

Subsequent studies have investigated BIM-assisted detailed estimating methodologies, demonstrating significant improvements in visualisation, quantity take-off accuracy and estimation efficiency. Shen and Issa (2010) quantitatively evaluated BIM-assisted detailed estimating (BADE) tools, concluding that BIM-assisted workflows significantly outperform traditional manual

estimation methods, especially for complex projects. Nevertheless, the authors identified several limitations, including insufficient knowledge of contextual construction and the inability of BIM systems to interpret non-standardised textual construction descriptions directly. Subsequently, several researchers explored ontology-based and semantic approaches to improve interoperability and standardisation in cost estimation systems. For example, Abanda et al. (2017) proposed a BIM-based ontology using the New Rules of Measurement (NRM) to enhance the consistency and machine readability of construction cost estimation. The study highlighted that conventional cost estimation is still 'manual, time-consuming and error-prone', and demonstrated the potential of semantic web technologies to improve BIM-based quantity take-off workflows. However, ontology-driven approaches still rely heavily on structured BIM environments and predefined classification systems.

More recent studies have introduced Natural Language Processing (NLP) and machine learning methods to process textual construction documents. Akanbi and Zhang (2020, 2021) proposed an NLP-based framework for extracting design information from construction specifications in order to support automated cost estimation. Their findings showed that extracting information from these specifications could substantially reduce manual effort and estimation errors while achieving high precision and recall rates. Nevertheless, the proposed approach focused primarily on extracting supplementary specification information, rather than addressing the broader challenge of semantically restructuring heterogeneous, spreadsheet-based cost estimation datasets.

Similarly, Deza et al. (2022) proposed machine learning models to automatically classify construction cost documents according to the International Construction Measurement Standard (ICMS). They emphasised that construction cost documentation remains highly heterogeneous and unstandardised, with a strong dependence on free-text descriptions. Reported classification accuracy exceeding 90% demonstrated the viability of semantic machine learning approaches for interpreting construction cost documents. However, the study primarily focused on text classification and standardisation benchmarking rather than automated estimate restructuring, pricing recommendations or enterprise-level integration.

There have also been parallel developments in sustainability-oriented BIM applications. Sun and Park (2020) developed a method of calculating CO_2 emissions during construction processes using BIM, integrating material quantities and environmental indicators into BIM workflows. Their study showed that BIM can support the evaluation of environmental performance and carbon-oriented construction analysis. However, existing sustainability-oriented approaches typically assume structured BIM datasets and do not address the challenge of extracting environmental intelligence from diverse historical estimation repositories.

These studies collectively demonstrate substantial progress in BIM-assisted estimating, ontology engineering, NLP-based information extraction and sustainability-oriented construction analysis. However, several research gaps remain that have not been addressed sufficiently:

a) Heterogeneous spreadsheet-based estimates remain difficult to standardise automatically;

b) existing BIM-centric approaches depend heavily on predefined model structures and classification systems;

c) semantic similarity matching between historical and newly uploaded estimates remains underexplored;

d) integration of pricing intelligence with sustainability-oriented CO_2 enrichment is still limited;

e) enterprise-level transformation of arbitrary XLSX estimates into machine-readable structured datasets has not been comprehensively addressed using LLM-based architectures.

The present study addresses aforementioned limitations by proposing an integrated LLM-based framework capable of:

1) Recognising heterogeneous construction estimation spreadsheets;

2) semantically restructuring estimate entities into standardised machine-readable formats;

3) retrieving historical pricing intelligence using semantic similarity matching;

4) enriching estimate positions with associated CO_2 indicators;

5) continuously improving recognition performance through reinforcement-learning feedback mechanisms.

Unlike previous BIM-centric or ontology-only approaches, the proposed framework focuses specifically on arbitrary, spreadsheet-based procurement and estimation workflows, which remain prevalent in the practical operations of the construction industry. Consequently, this study advances AI-assisted cost estimation by combining semantic document understanding, historical knowledge retrieval and sustainability-oriented enrichment within a unified, enterprise-level automation framework.

3. Methods and Procedures

The proposed framework incorporates a multi-stage AI pipeline specifically designed for processing heterogeneous construction cost estimation spreadsheets. The methodology consists of four main stages: dataset preparation; semantic annotation; LLM-based recognition; and historical similarity retrieval.

Initially, a large collection of historical Excel-based construction estimates was gathered from completed projects. These documents exhibited significant

variability in terms of formatting, terminology, structure and linguistic expression. In order to prepare the dataset for model training, the positions, quantities, units, pricing attributes and structural relationships of the estimates were manually annotated and normalised into a unified schema.

The annotated datasets were then used to fine-tune the LLM recognition pipeline. The framework uses automated spreadsheet parsing mechanisms combined with semantic interpretation models that can recognise estimated positions independently of spreadsheet layout or terminology variations. The system architecture supports asynchronous processing via API-based integration, enabling uploaded spreadsheets to be automatically converted into structured database entities.

An additional semantic similarity retrieval module was implemented to compare newly uploaded estimated positions with historical datasets. Despite linguistic variations, vector-based similarity matching techniques were applied to identify semantically related construction positions. For instance, estimates such as 'Entrance Gate Construction', 'Fence Gate Installation', and 'Pedestrian Access Gate Assembly' could be recognised as semantically equivalent, with similarity scores exceeding 85%.

The structured outputs generated by the framework are stored in enterprise databases. These outputs include information on estimated positions, quantities, unit costs, total costs and associated sustainability indicators. For selected construction materials, such as concrete, steel and cement products, historical CO_2 emission data has also been linked to recognised estimate positions to support sustainability-oriented assessments.

a) Multi-stage Automation Pipeline

As illustrated in Figures 1 and 2, the automation workflow begins with the upload of an XLSX file or folder containing one or more construction cost estimate documents. These documents may include bills of quantities, procurement specifications, contractor quotations or historical priced estimates. As such files often have different layouts and terminology, the system first extracts the spreadsheet content and metadata. This includes worksheet names, cell values, table boundaries, row structures, column headers, merged cells, units, quantities and cost-related fields.

After extraction, the system automatically creates a source profile. This profile determines how the spreadsheet is interpreted, identifying whether it contains historical prices, an empty tender estimate, material quantities or multiple estimate sections, for example. The source profile is represented as:

$$Ps = f(M,T,H,C), \quad (1)$$

where Ps is the source profile, M represents metadata, T represents detected tables, H represents headers, C represents cell-level content.

The subsequent stage of the process is the parsing of worksheets and tables. This involves the separation of relevant estimate tables from non-relevant spreadsheet content and labels. The estimate entities to be identified include position descriptions, units, quantities, unit costs and total costs. In a simplified form, each estimate item is structured as:

$$e_i = \{d_i, u_i, q_i, c_i, t_i\}, \quad (2)$$

where d_i is the work-item description, u_i is the unit, q_i is the quantity, c_i is the unit cost, and t_i is the total cost.

The labelled entities are then processed by the LLM-based restructuring module. This module uses a domain-specific RIXBUILD pre-trained transformer model, as well as external AI APIs if necessary, to normalise terminology, resolve ambiguous descriptions, translate or paraphrase text and convert spreadsheet data into standardised, machine-readable JSON or XML structures. The cost of each recognised item may be checked using:

$$t_i = q_i \times c_i, \quad (3)$$

where t_i is the total cost of item i, q_i is its quantity, and c_i is its unit cost.

The structured estimate data are stored in the enterprise database. Once stored, each newly recognised estimate position is compared with historical estimate records. Semantic similarity matching is performed by embedding the new item description and historical item descriptions into a vector space. A common similarity measure is a cosine similarity:

$$S(d_i, d_j) = \frac{\vec{d}_i \cdot \vec{d}_j}{\|\vec{d}_i\| \|\vec{d}_j\|}, \quad (4)$$

where $S(d_i,d_j)$ denotes the similarity between the new description d_i and a historical description d_j. If the similarity score is equal to or greater than a predefined threshold, the system retrieves historical unit prices and recommends them for the new estimate position.

When several historical matches are available, a weighted historical price can be calculated as:

$$\hat{c}_i = \frac{\sum_{j=1}^{n} S_{ij} c_j}{\sum_{j=1}^{n} S_{ij}}, \quad (5)$$

where $\hat{c}_i$ is the recommended unit cost for the new item, S_{ij} is the similarity score between the new item and historical item j, and c_j is the historical unit cost.

If CO_2 data are available for the matched historical item, the system attaches sustainability indicators to the new estimate position. The estimated CO_2 impact can be calculated as:

$$CO_{2,i} = q_i \times EF_i, \quad (6)$$

where CO_2 is the estimated carbon impact of item i, q_i is the quantity, and EF_i is the emission factor per unit. If the similarity score is below the threshold, or if required CO_2 data are missing, the item is flagged

for manual review. This ensures that outputs that are uncertain or incomplete are not accepted automatically without validation by the estimator.

The quality check stage verifies the consistency of the following: extracted entities; calculated costs; similarity-based recommendations; and CO_2 enrichment. If the output is accepted, the validated estimate is saved and added to the historical dataset. If the output is rejected, the estimator corrects the result manually, and this correction is stored as feedback for improving future models.

Finally, the validated, structured estimate is exported to systems such as cost management, procurement, ERP, BIM and sustainability reporting. This completes the automation loop, transforming heterogeneous, spreadsheet-based estimate documents into standardised, reusable enterprise data that is enriched for sustainability.

b. Framework Deployment Architecture

In order to evaluate the practical applicability of the proposed framework within real-world procurement and cost estimation workflows, the developed solution was deployed as an enterprise-oriented, cloud-integrated microservice architecture. This deployment environment was designed to support the scalable processing of heterogeneous XLSX construction estimates, as well as asynchronous API communication, semantic retrieval operations and integration with enterprise-level cost management systems.

The deployment architecture consists of five primary layers:

a) Bocument ingestion and preprocessing layer;
b) semantic parsing and LLM inference layer;
c) vector similarity retrieval layer;
d) sustainability enrichment and validation layer;
e) enterprise integration and persistence layer.

The ingestion layer accepts XLSX files or folders, which can be uploaded either manually via a web interface or automatically via API endpoints. The uploaded documents are transferred to a pre-processing queue, where spreadsheet normalisation, metadata extraction, worksheet segmentation and table boundary detection are performed.

The semantic inference layer comprises the internally trained RIXBUILD transformer model, as well as external LLM APIs that are used for contextual interpretation, semantic restructuring and terminology normalisation. Deployment supports hybrid inference execution, prioritising locally hosted, domain-specific models for sensitive construction datasets, while selectively using external APIs for complex contextual interpretation tasks.

The vector retrieval subsystem uses embedding-based semantic indexing to match historical estimates. Historical estimate positions are converted into vector embeddings and stored in an optimised vector database for approximate nearest-neighbour retrieval. The semantic retrieval process can be formally expressed as:

$$R(q) = \arg\max_{d_j \in D} S(q, d_j), \quad (7)$$

where $R(q)$ denotes the retrieved historical estimate item for query position q, while D represents the historical estimate dataset, and $S(q,d_j)$ is the semantic similarity score.

To improve enterprise interoperability, all recognised estimation entities are exported using standardised JSON and XML schemas that are compatible with standard ERP systems, BIM-integrated workflows, procurement systems, sustainability assessment platforms, and cost management databases.

The deployment also incorporates asynchronous job scheduling and distributed processing mechanisms to support the simultaneous processing of estimates by multiple users. An experimental deployment demonstrated stable operation when processing multiple XLSX estimate datasets, each with several thousand estimate positions, concurrently.

c. Dataset Preparation and Annotation Strategy

The quality of semantic recognition within the proposed framework depends heavily on the consistency and representativeness of the training datasets. Therefore, a dedicated annotation and standardisation strategy was developed to prepare historical estimate repositories. Approximately 100 proprietary historical construction estimates, containing validated pricing data, were collected from completed procurement and construction projects.

The uploaded spreadsheets exhibited substantial variability in terms of formatting, worksheet hierarchy, terminology, languages and language expression, representation of measurement units, pricing structures and currencies, among other factors. To improve semantic consistency, estimated values were manually normalised into standardised conceptual entities. The annotation procedures encompassed the semantic grouping of equivalent construction activities, the normalisation of units of measurement, the removal of duplicate expressions, the correction of inconsistent abbreviations, and the mapping of synonymous construction terminology.

The semantic annotation process generated structured training pairs of the following form:

$$Ai = (d_i, c_i, u_i, q_i, CO_{2,i}) \quad (8)$$

where d_i represents the normalised description, c_i the historical unit cost, u_i the unit of measurement, q_i the quantity information, and $CO_{2,i}$ is the associated sustainability indicator. The utilisation of such structured representation facilitated the framework's capacity to discern semantic relationships between linguistically disparate yet operationally analogous construction estimate positions.

d. Evaluation Methodology

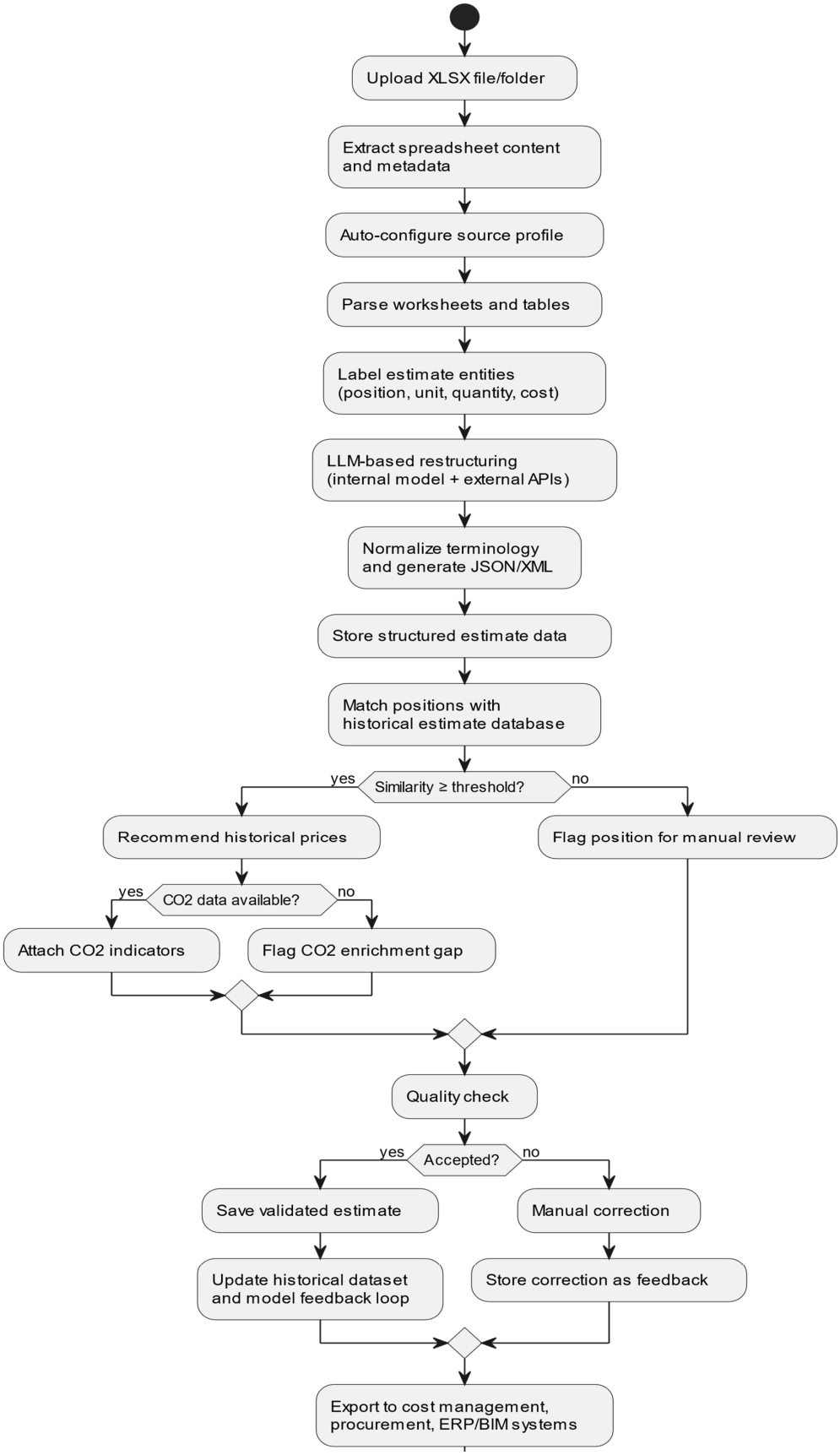


Figure 1. The proposed LLM-Based Cost Estimate Recognition and Enrichment Workflow

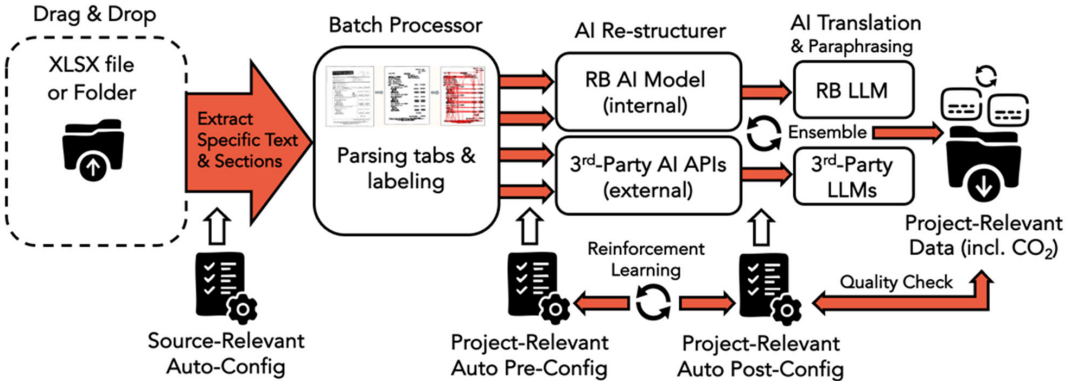


Figure 2. The overview of the main automation processes within the proposed RIXBUILD Framework

The evaluation of the proposed framework was performed using quantitative and operational performance metrics designed to assess the following: (a) semantic recognition accuracy, (b) estimate completion efficiency, (c) pricing recommendation reliability, (d) sustainability enrichment applicability, and (e) operational workload reduction. The experimental evaluation was conducted on approximately 100 historical construction estimate spreadsheets. These spreadsheets contained validated pricing information and manually verified estimate structures.

4. Results

The experimental validation demonstrated that the proposed framework successfully recognised and structured heterogeneous construction estimate spreadsheets with high semantic accuracy. The semantic similarity retrieval mechanism achieved position matching accuracy exceeding 85% for semantically related estimate entries.

The testing process involved the utilisation of approximately 100 historical construction estimates, each comprising validated pricing information. The results indicated that the framework had the capacity to automatically pre-fill approximately 50-70% of newly uploaded 'empty' estimates based on historical similarity retrieval. Consequently, estimators were provided with automatically suggested unit prices and related historical references for semantically matched positions.

The proposed approach resulted in a substantial reduction in manual workload during procurement preparation and tender response generation. In comparison with conventional manual estimation procedures, the framework reduced the effort required for estimate preparation by approximately 70%. The most substantial efficiency gains were observed in repetitive procurement scenarios involving historically recurring construction positions.

Furthermore, the integration of sustainability-related metadata enabled automatic approximation of CO_2 indicators for newly recognised estimate positions. This functionality provides an additional decision-support layer for environmentally oriented procurement and sustainable construction planning.

Experimental observations also suggest that the framework's effectiveness increases proportionally with the size of the historical estimate repository, the semantic consistency of the annotated datasets, and the diversity of the previously processed procurement scenarios.

Consequently, continuous dataset standardisation and reinforcement-learning feedback mechanisms are critical for improving the long-term performance of the framework.

5. Conclusions

This study presents a framework based on LLMs for the automated recognition, semantic structuring and enrichment of construction cost estimates from heterogeneous spreadsheet documents. The results obtained demonstrate the practical applicability of large language models for improving the operational efficiency of construction procurement and cost management processes.

The main conclusions of the research are as follows: Firstly, the proposed framework successfully converts heterogeneous Excel-based estimates into structured, machine-readable datasets. Secondly, semantic similarity retrieval enables the effective matching of construction positions despite substantial linguistic variability. Thirdly, reuse of historical estimates significantly accelerates preparation of new procurement estimates. Furthermore, the developed solution reduces the effort required for manual estimation by around 70%, compared to the conventional workflows currently used by the construction industry. Finally, integrating CO_2-related metadata introduces additional sustainability-oriented analytical capabilities. It is worth noting that

LLM-assisted recognition frameworks demonstrate strong potential for broader application within digital construction management ecosystems, and further in-depth studies are certainly required.

Future research should focus on expanding multilingual recognition capabilities, improving semantic precision for highly specialised construction domains, and integrating Building Information Modeling (BIM) datasets with cost estimation pipelines.

Acknowledgment

This research has been supported by the European Union Recovery and Resilience Facility through the Latvian Recovery and Resilience Plan within the project "Development of an artificial intelligence solution for automated identification, structuring and analysis of construction cost estimate data to facilitate the calculation of CO_2 emissions at different stages of construction". Project No. 2.2.1.3.i.0/2/24/A/CFLA/001.

References:

Abanda, F. H., Kamsu-Foguem, B., & Tah, J. H. M. (2017). BIM – New rules of measurement ontology for construction cost estimation. *Engineering Science and Technology, an International Journal*, 20(2), 443–459. https://doi.org/10.1016/j.jestch.2017.01.007

Akanbi, T., & Zhang, J. (2020). Automated design information extraction from construction specifications to support wood construction cost estimation. In Construction Research Congress 2020 (pp. 658-666). American Society of Civil Engineers. https://doi.org/10.1061/9780784482889.069

Akanbi, T., & Zhang, J. (2021). Design information extraction from construction specifications to support cost estimation. *Automation in Construction*, 131, 103835. https://doi.org/10.1016/j.autcon.2021.103835

Deza, J. I., Ihshaish, H., & Mahdjoubi, L. (2022). A machine learning approach to classifying construction cost documents into the International Construction Measurement Standard. arXiv. https://doi.org/10.48550/arXiv.2211.07705

Elmousalami, H., Maxy, M., Hui, F. K. P., & Aye, L. (2025). AI in automated sustainable construction engineering management. *Automation in Construction*, 175, 106202. https://doi.org/10.1016/j.autcon.2025.106202

Jadid, M. N., & Idrees, M. M. (2007). Cost estimation of structural skeleton using an interactive automation algorithm: A conceptual approach. *Automation in Construction*, 16(6), 797–805.

Shen, Z., & Issa, R. R. (2010). Quantitative evaluation of the BIM-assisted construction detailed cost estimates. https://digitalcommons.unl.edu/constructionmgmt/4/

Sun, H., & Park, Y. (2020). CO2 emission calculation method during construction process for developing BIM-based performance evaluation system. *Applied Sciences*, 10(16), 5587. https://doi.org/10.3390/app10165587