Joint CSCE Construction Specialty & CRC Conference 2025
*Conférence conjointe spécialisée en construction de la SCGC et CRC-2025*

Montreal, Quebec
July 28-31, 2025 / *28-31 juillet 2025*

# Automated Sustainable Evaluation for Construction Contracts Using Machine Learning

Mohammad Sabek[1], Baseel Andres Ammar[1], Ahmed Hammad[1], Farook Hamzeh[1] and Vicente Gonzalez[1]

[1] Department of Civil and Environmental Engineering, University of Alberta, Edmonton, Canada

**ABSTRACT:** The construction industry is a major contributor to environmental degradation, resource consumption, and greenhouse gas emissions. This research proposes an innovative, automated system for the sustainable evaluation of construction contracts by integrating advanced machine learning (ML) techniques and large language models (LLMs). The proposed framework automates extracting, classifying, and evaluating sustainability-related clauses from diverse contract documents. The system transforms contractual text into semantic embeddings stored in a dedicated vector database by leveraging Natural Language Processing (NLP) tools such as LangChain for text segmentation and FAISS for efficient vector retrieval . A locally deployed LLM then processes these embeddings to identify key sustainability factors, including energy efficiency, waste management, environmental impact, and worker safety, providing comprehensive insights in real time. We tested the system on a set of construction contracts, which demonstrated the system's capability to accurately classify sustainability clauses into environmental, social, and economic categories. This capability contributes to a reduction in human error and a significant enhancement in the efficiency of the evaluation process. The results underscore the potential of ML-driven approaches to standardize and enhance sustainability assessments during the contract formulation stage, a critical yet under-explored phase in construction management. Moreover, integrating keyword-scoring mechanisms further refines the system's analytical precision, ensuring that sustainability objectives are clearly defined and met. This research contributes to the development of data-driven sustainable procurement practices and sets the stage for future enhancements, including deeper integration with external sustainability benchmarks and broader industry adoption.

## 1. INTRODUCTION

The construction industry plays a pivotal role in economic development, yet it is also one of the most significant contributors to environmental degradation, resource depletion, and greenhouse gas emissions (Escamilla et al., 2016). The sector is responsible for approximately 35% of global $CO_2$ emissions and generates 45% to 60% of landfill waste (Escamilla et al., 2016). Given the increasing emphasis on sustainable development, it is essential to integrate sustainability considerations into construction projects from the early stages, particularly within contractual agreements. Contracts are legally binding documents defining project requirements, responsibilities, and performance expectations. However, existing sustainability assessment frameworks primarily focus on pre- and post-construction phases, often overlooking sustainability clauses during contract formulation and execution (Illankoon et al., 2017). This gap in early sustainable evaluation (ESE) limits the effectiveness of sustainability measures throughout the project lifecycle.

This study proposes an automated sustainability evaluation system for construction contracts using Machine Learning (ML) and Large Language Models (LLMs). Traditional contract analysis methods rely heavily on manual review, which is time-consuming, prone to human error, and lacks standardization (Hasan et al., 2021). This research aims to automate the extraction, classification, and evaluation of sustainability-related clauses in construction contracts by incorporating components like LLM, ML, and Natural Language Processing (NLP). The system will identify critical sustainability factors, including energy efficiency, waste management, environmental impact, and worker safety, enabling a more data-driven and standardized approach to contract assessment.

## 2. LITERATURE REVIEW

### 2.1 Sustainability in Construction

The construction industry has paid considerable attention to sustainability (Spence et al., 1995). According to research findings, the construction industry is responsible for 35% of $CO_2$ emissions and produces 45% to 60% of waste that ends up in landfills (Escamilla et al., 2016). Numerous rating systems, such as LEED, CASBEE, and BREEAM, have been developed by various organizations worldwide. (Dubey et al., 2023). These systems aim to validate construction projects as green projects, primarily evaluating their environmental, social, and economic effects throughout the project's entire lifecycle, from initiation to demolition (Dubey et al., 2023). However, in many of these rating systems, the parameters concerning sustainability during the construction phase have been confined to issues such as soil erosion and the responsible use of materials. Other vital facets of construction sustainability, such as energy consumption, noise control, worker safety, and health, as well as the impact on the surrounding environment, have not been adequately considered (Illankoon et al., 2017).

These rating systems predominantly address the parameters before and after the construction phases. Regrettably, many of these systems have not provided sufficient consideration to various factors during the built environment stage, with a few exceptions where construction was a prerequisite for certification. Nonetheless, it is essential to carry out sustainability evaluations before the commencement of the construction contract phase to guarantee a thorough assessment of a wide range of construction projects right from the outset and maximize sustainability benefits.

### 2.2 Content Analysis of Contracts

A methodical approach was introduced in 1989 grounded in an impartial analysis. This technique facilitates examining extensive written content by considering data frequencies, meanings, and interrelationships (Krippendorff, 1989). Incorporating sustainability within construction contracts has been the subject of substantial research. Academics have stressed the significance of including sustainability clauses and key performance indicators (KPIs) to ascertain that sustainability goals were articulated and quantifiable (Lam, 2022). Researchers had concluded that the review of empirical studies illustrated that a comprehensive assessment of sustainability in development projects can be gauged through project outcome Key Performance Indicators (KPIs) (Lam, 2022). Economic sustainability can be evaluated using cost-related KPIs, while quality-related KPIs can serve as metrics for assessing social and environmental sustainability. The results of a quantitative hierarchical regression analysis substantiated that these three dimensions of sustainability are notably influenced by task performance and contextual performance factors. This alignment with job performance theory strengthened this association.

## 2.3 Sustainability Inclusion in Construction Contracts Index

In a broad sense, contracts played a crucial role in safeguarding the financial interests of involved parties and ensuring that neither parties incurred losses; The clear articulation of rights and obligations during the early stages of a project played a critical role in ensuring contractual effectiveness and minimizing disputes (Mitkus & Trinkūnienė, 2008). As a result, well-structured contracts also promote social and environmental sustainability by reinforcing economic stability throughout the project lifecycle. Furthermore, many clauses can be introduced into the construction contract to promote favorable sustainability results. Incorporating constructability considerations and construction methods offered a vital advantage in progressing toward a sustainable construction sector (Babaeian Jelodar et al.,2013).

The importance of sustainability in construction contracts has led to the inclusion of specific clauses related to environmental, social, and economic objectives in standardized contract templates. (Hibberd, 2017). Standard contractual terms may not comprehensively encompass all three aspects of sustainability, nor can they fully grasp the specific construction procedures and sequences. As a result, some scholars have explored the integration of sustainability into construction contracts through dedicated sustainability clauses. Embracing sustainability practices can prove highly advantageous for contractors, as it may be viewed as a competitive edge. Sustainability has become a criterion for contractor selection primarily because of the advantages it can give during the bidding phase (Manoliadis and Vatalis, 2016).

A comprehensive study on sustainable practices in construction contracts was conducted (Hasan et al., 2021). The study highlighted the need for clear, well-defined sustainability clauses covering various aspects of sustainability, such as energy efficiency, waste management, and social responsibility. It underlined the significance of using contracts to foster sustainability within the construction industry (Hasan et al., 2021).

## 2.4 Machine Learning in Legal Text Analysis

Machine learning methods have contributed significantly to the analysis of real-world data in support of decision-making processes (Kaab et al., 2019). A novel framework was proposed using supervised learning techniques, specifically Classification and Regression Trees (CART) and Adaptive Neuro-Fuzzy Inference System (ANFIS), to evaluate sustainability performance (Mehrbakhsh et al., 2019).

Machine learning techniques, particularly NLP, have made significant inroads into legal text analysis. NLP and machine learning algorithms have been deployed for various contract analysis and legal text mining aspects. Automating contract analysis is a promising avenue, including extracting and categorizing sustainability-related clauses. The mentioned studies demonstrated the potential of machine learning to identify and categorize contract terms and clauses.

## 2.5 Machine Learning and NLP Applications in Construction

Integrating machine learning and natural language processing techniques has demonstrated significant potential in contract management, revealing new dimensions in risk identification and mitigation. One study applied ML to identify and manage risks embedded in inter-organizational contracts, highlighting the effectiveness of advanced algorithms in analyzing complex contractual agreements and offering valuable insights for organizations involved in collaborative ventures (Ariai et al., 2024).

Another study extended the application of ML to airport project contracts, emphasizing the efficient detection of major changes in contract documents. By applying NLP techniques, the researchers improved the change identification process, enabling stakeholders to respond promptly and maintain effective project control (Khalef & El-Adaway, 2021). Similarly, NLP methods were used to extract clauses that could pose risks to contracting parties, contributing to the interpretation of critical contract elements and supporting a more informed approach to negotiation and risk mitigation (Lee et al., 2019).

Considering the evolving emphasis on sustainability in the construction industry, this study seeks to address the following key research question: How can Early Sustainable Evaluation (ESE) be effectively integrated into the assessment framework of construction contracts, and to what extent can machine learning contribute to the early identification and quantification of sustainability factors. By exploring these questions, our research aims to advance the understanding of sustainable practices in contract evaluation and provide valuable insights for enhancing the early identification and management of sustainability considerations within construction projects.

## 3. METHODOLOGY

In this section, we are discussing how we have built our evaluation system, which has many components that will be addressed simply due to the size limitation of this conference paper. An overview of the system components can be seen in Figure 1:

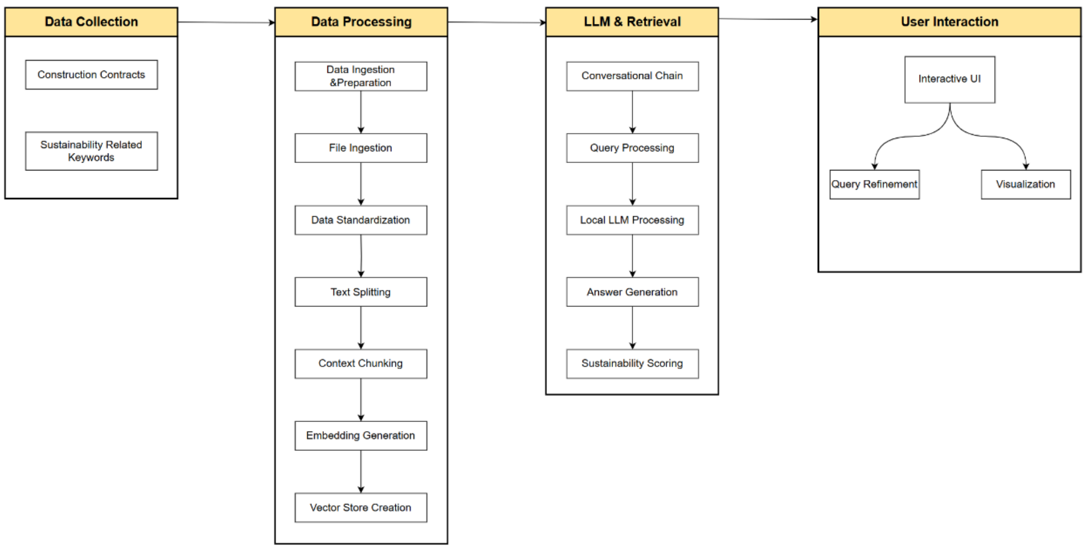


Figure 1: Structured Approach to Building the Evaluation System

### 3.1 Data Collection

The workflow begins with the collection of contract documents in PDF, DOCX, or TXT formats. In parallel, an extensive search was performed to collect keywords related to sustainability clauses from academic literature and practical case studies. A framework of some of the keywords searched for in relation to sustainability clauses and contracts is presented in Figure 2.

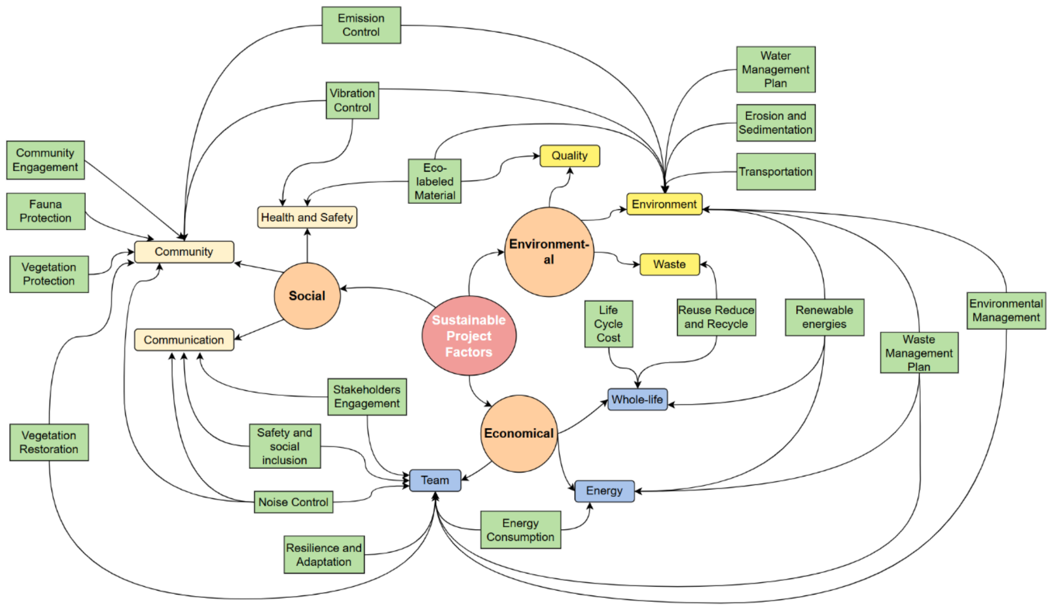


Figure 2: Sustainable project factors

### 3.2 Data Processing

The uploaded documents are first loaded into the system using a custom-built ingestion pipeline that is engineered to process and normalize various file formats by employing format-specific parsing algorithms. To accurately identify and extract content, the pipeline leverages libraries such as Apache Tika for file format detection, Python-docx for processing DOCX files, and PyPDF2 for PDF extraction. These tools enable the system to recognize and accommodate the structural nuances and size variations inherent in contract documents, ensuring that the extracted text is standardized and optimized for subsequent segmentation and analysis. This step establishes a standardized input pipeline, ensuring subsequent text splitting and embedding processes receive clean, consistently formatted data.

Moreover, during the ingestion process, the system extracts metadata, including a unique contract identifier (ID) from each document (if available) or assigns one automatically. This unique ID is maintained throughout the processing pipeline, ensuring that every extracted clause can be accurately associated with its original contract for precise document-level analysis. In parallel, the system references the established set of sustainability-related keywords.

In this implementation, the default LangChain splitting algorithm was customized to suit the nuances of contractual language better. This modification ensures that legal clauses are preserved in context, which is crucial for accurately identifying sustainability-related content. By working with smaller units of text, the system can more accurately locate relevant clauses and identify potential sustainability-related issues, laying the groundwork for precise retrieval and analysis.

The embedding process was fine-tuned by incorporating domain-specific modifications. The integration of FAISS and ChromaDB was optimized to accelerate retrieval speed and improve accuracy in matching contract clauses, thereby tailoring the process to the specific demands of contract analysis. Concurrently, user queries (e.g., requests for specific sustainability clauses) are also converted into embeddings for subsequent matching. The system can incorporate FAISS as the engine for efficient nearest-neighbor searches, ensuring the rapid retrieval of relevant text chunks from the ChromaDB vector store (Douze et al., 2024). This architecture enables the processing of large volumes of embedded data with minimal computational overhead, allowing for real-time or near-real-time query responses.

### 3.3 LLM and Conversational Chain Retrieval

Once the vector store is established, the system constructs a conversational chain to facilitate interactive queries about the uploaded contracts. Our developed conversational chain mechanism adapts iteratively in response to user interactions. This custom design enhances semantic matching between user queries and contract segments, ensuring that the system delivers more context-aware and relevant responses. When a user asks, “Which clauses address waste management? “The system embeds this query and searches the ChromaDB vector store for the most similar contract chunks (ChromaDB, 2023). This retrieval step leverages the embeddings generated earlier, matching semantic content rather than relying solely on keyword overlap. In doing so, the system pinpoints the contract sections most relevant to the user’s inquiry, even if the language used does not match the user’s query.

With the relevant text chunks identified, the system engages a locally deployed large language model (LLM); in our case, we have used the IBM Granite3.1-MoE 3B model, which is a long-context mixture of experts (MoE) Granite model from IBM designed for low-latency usage (IBM, 2023). Granite3.1 doesn’t require a high-performance GPU to run it via Ollama due to its small size (2.5GB). Ollama is a tool that allows running LLM locally on computers without internet access to running the LLM (Ollama,2023). Granite3.1-MoE 3B model performs multiple functions: it interprets the user’s question, synthesizes the retrieved contract segments, and integrates any relevant keyword-based sustainability scoring that has been computed. The local deployment addresses privacy concerns inherent in contract analysis and reduces latency, ensuring a more responsive user experience. In addition to answering direct user questions, the LLM can generate or refine contract clauses, highlighting ways to embed environmental, social, or economic sustainability features into the document.

### 3.4 Interactive User Experience and Iterative Querying

The system accommodates iterative user interactions. After receiving an initial answer, users may refine or expand their queries, for instance, by asking how risk allocation might be balanced alongside community engagement or how performance evaluation aligns with the identified sustainability factors. The system then re-embeds the new question, searches the vector store again, and re-engages the LLM logic, thus providing updated responses. This cyclical, conversational flow allows stakeholders to explore contract documents comprehensively, investigating multiple facets of sustainability in a dynamic, question-and-answer format. The user interface was developed using Streamlit, an open-source Python framework for data scientists and AI/ML engineers to deliver interactive data apps (Streamlit ,2023), which can be seen in Figure 3.

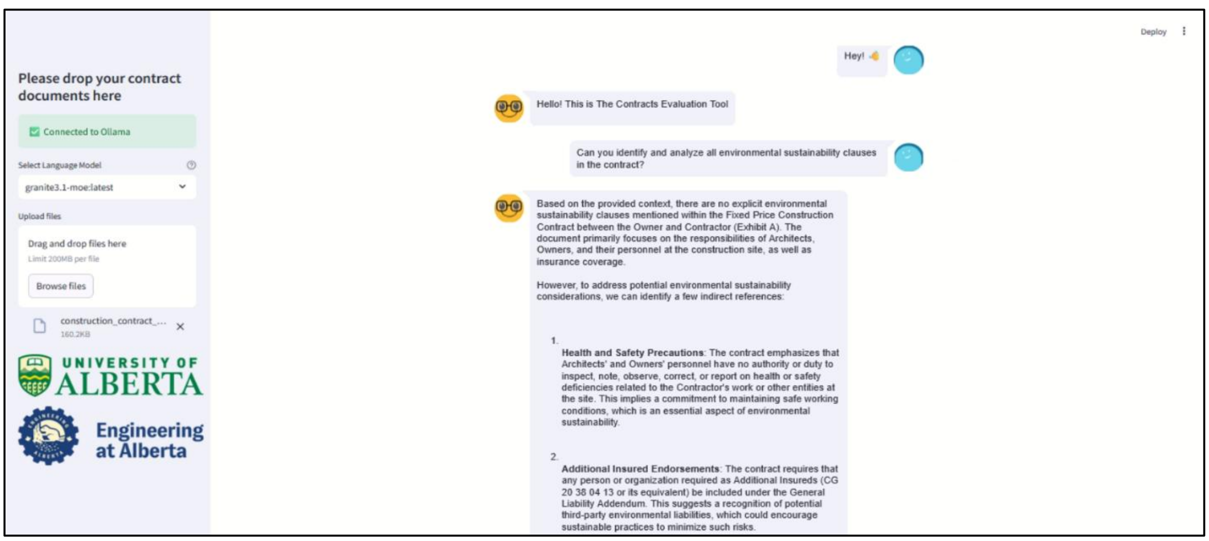


Figure 3: Interactive user interface using Streamlit

In addition to the intuitive graphical user interface built with Streamlit, our system incorporates multithreading in the backend to enhance performance. Specifically, up to 8 contract files can be processed concurrently, which not only speeds up the evaluation process but also ensures that the system remains within the context limitations of the deployed LLM. This concurrent processing capability significantly enhances scalability and efficiency, making the tool well-suited for handling large volumes of contracts in real-time.

### 3.5 Alignment with Sustainability Keyword Analysis

Our methodology includes a sustainability keyword-scoring mechanism that operates in parallel. Before or during LLM-based analysis, contract chunks can be scored according to the presence or absence of key sustainability terms (e.g., “carbon footprint,” “community engagement,” “life-cycle cost”). These scores may inform the LLM’s reasoning, prompting it to recommend contract revisions or enhancements where certain sustainability aspects are found to be deficient. Thus, the synergy between chunk-level embedding, semantic retrieval, and keyword scoring ensures that contextual depth and sustainability objectives remain central to the system’s outputs.

## 4. TOOL VALIDATION

The tool’s performance was meticulously validated using a curated dataset comprising ten distinct construction contracts that varied in format and complexity. Each contract was processed through the system’s automated pipeline, from initial document ingestion to semantic embedding and LLM-based clause analysis. Expert reviewers then compared the tool’s output with manual assessments, focusing on accurately extracting and categorizing sustainability clauses. This evaluation confirmed that the system consistently identified and classified key sustainability clauses across environmental, social, and economic dimensions.

The validation process highlighted the tool’s robustness in handling heterogeneous data formats and its capacity to preserve contextual relevance during text segmentation and semantic retrieval. While the current validation demonstrates promising performance, it also revealed areas for further enhancement, such as refining the classification algorithm for more complex contractual language. These insights provide a clear roadmap for iterative improvements and underscore the tool’s potential as a reliable decision-support resource for sustainable construction project management. To quantitatively evaluate our system’s performance, we conducted experiments on a dataset of ten diverse construction contracts. Table 1 summarizes the key metrics obtained during validation, demonstrating the system’s effectiveness in accurately identifying and extracting sustainability-related clauses. These metrics underscore the reliability and efficiency of our automated evaluation process

Table 1: Quantitative Performance Metrics for Sustainability Clause Extraction

| Metric | Value | Description |
|---|---|---|
| Overall Accuracy | 92% | Percentage of sustainability clauses correctly identified compared to manual annotations. |
| Precision | 90% | Proportion of identified clauses that are truly relevant (i.e., correctly extracted sustainability clauses). |
| Recall | 88% | Proportion of actual sustainability clauses that were successfully identified by the system. |
| Average Processing Time | 2.3 sec / 1 page<br>Using an Nvidia 4090 RTX GPU | Average time taken to process and analyze each contract file. |

## 5. CONCLUSION

In conclusion, developing an automated, sustainable evaluation system for construction contracts significantly advances the integration of environmental, social, and economic sustainability into the early stages of project planning. The research addresses critical gaps in traditional manual assessments by harnessing state-of-the-art ML techniques and LLMs, offering a robust framework that enhances accuracy, efficiency, and consistency. The system's ability to parse complex contractual language and extract pertinent sustainability clauses paves the way for more informed decision-making in the construction industry, ultimately promoting greener and more responsible project practices. While initial validation efforts have demonstrated the tool's potential, the findings also indicate the need for continued refinement and broader testing. This study highlights the transformative impact of digital technologies on sustainable contract management, establishing a solid foundation for future research and development in this domain. It is important to note that our system is distinctly tailored for the automated evaluation of construction contracts, setting it apart from general-purpose LLM platforms such as ChatGPT.

While ChatGPT relies on manually designed prompts for general conversation, our tool employs a dedicated ingestion pipeline, domain-specific text splitting, and embedding techniques, as well as a robust keyword-scoring mechanism to identify and classify sustainability-related clauses. Moreover, by leveraging a locally deployed and fine-tuned IBM Granite3.1-MoE 3B model, our system addresses critical privacy and latency issues, and its multithreading capabilities enable the simultaneous processing of up to 8 contracts. These specialized features not only improve accuracy and efficiency but also provide a more structured and context-aware evaluation that is specifically designed for the complexities of construction contracts.

## 6. RECOMMENDATIONS

Building on the promising outcomes of this research, we recommend that future work integrate advanced NLP techniques such as Named Entity Recognition and sentiment analysis to further refine the extraction and evaluation of contractual sustainability nuances. Expanding the dataset to include a broader variety of contracts and larger sample sizes is essential to enhance predictive accuracy and validate system performance across different project types and regions. In addition, incorporating external sustainability databases and established certification frameworks (e.g., LEED) will enrich the evaluative context, enabling more comprehensive analyses of contract compliance.

Enhancing visualization and reporting features is also crucial for translating complex data into actionable insights and fostering stakeholder engagement. Furthermore, future iterations of the system will explore the potential of advanced thinking models, including fine-tuned R1 models and emerging solutions such as Llama 4 Scout which features 17B active parameters distributed among 16 experts (with a total of 109B parameters), an industry-leading context length of 10 million tokens (Meta AI,2025), which can enhance the system's capacity to process and analyze complex contractual data. Finally, ensuring standardization and achieving interoperability with existing contract management systems will be key to driving widespread adoption within the construction industry, ultimately embedding sustainable practices at every stage of project development.

# REFERENCES


Ariai, F. & Demartini, G. 2024. Natural Language Processing for the Legal Domain: A Survey of Tasks, Datasets, Models, and Challenges. *ArXiv*, abs/2410.21306.

Babaeian Jelodar, M., Jaafar, M. S. & Yiu, T. W. 2013. In search of sustainability: Constructability application and contract management in Malaysian industrialized building systems. *Journal of Legal Affairs and Dispute Resolution in Engineering and Construction*, 5(4), 196–204. https://doi.org/10.1061/(ASCE)LA.1943-4170.0000121

ChromaDB. 2023. Chroma: The Open-Source Embedding Database [Software]. Retrieved from https://www.trychroma.com

Douze, M., Guzhva, A., Deng, C., Johnson, J., Szilvasy, G., Mazaré, P.-E., Lomeli, M., Hosseini, L. & Jégou, H. 2024. The FAISS library. *arXiv preprint* arXiv:2401.08281. Retrieved from https://faiss.ai/

Dubey, M. K., et al. 2023. Need for a rating system for assessing the sustainability of the built environment during the construction stage. *Total Environment Research Themes*, 7, 100061. https://doi.org/10.1016/j.totert.2023.100061

Escamilla, E. Z., Habert, G. & Wohlmuth, E. 2016. When $CO_2$ counts: Sustainability assessment of industrialized bamboo as an alternative for social housing programs in the Philippines. *Building and Environment*, 15(3), 223–239. https://doi.org/10.1080/014461997372971

Hasan Mahmoud & Beheiry, S. 2021. Sustainability Inclusion in Construction Contracts Index. *Journal of Legal Affairs and Dispute Resolution in Engineering and Construction*, 13(4). https://doi.org/10.1061/(ASCE)LA.1943-4170.0000503

Hibberd, P. 2017. Sustainability—The role of construction contracts. In *Future challenges in evaluating and managing sustainable development in the built environment*, 268–284. New York: Wiley. https://doi.org/10.1002/9781119190691.ch16

IBM. 2023. IBM Granite 3.1-MoE 3B Model [Internal model documentation]. Retrieved from https://www.ibm.com/granite/docs/models/granite/

Illankoon, I. M. C. S., Tam, V. W. Y., Le, K. N. & Shen, L. 2017. Key credit criteria among international green building rating tools. *Journal of Cleaner Production*, 164, 209–220. https://doi.org/10.1016/j.jclepro.2017.06.206

Kaab, A., Sharifi, M., Mobli, H., Nabavi-Pelesaraei, A. & Chau, K.-W. 2019. Combined life cycle assessment and artificial intelligence for prediction of output energy and environmental impacts of sugarcane production. *Science of The Total Environment*, 664, 1005–1019. https://doi.org/10.1016/j.scitotenv.2019.02.004

Khalef, R. & El-adaway, I.H. 2021. Automated Identification of Substantial Changes in Construction Projects of Airport Improvement Program: Machine Learning and Natural Language Processing Comparative Analysis. *Journal of Management in Engineering*, 37, 04021062.

Krippendorff, K. 1989. Content analysis. In E. Barnouw, G. Gerbner, W. Schramm, T. L. Worth & L. Gross (Eds.), *International encyclopedia of communication* (Vol. 1, pp. 403–407). Malden, MA: Blackwell.

Lam, T. Y. M. 2022. Driving sustainable construction development through post-contract key performance indicators and drivers. *Smart and Sustainable Built Environment*, 11(3), 483–499. https://doi.org/10.1108/SASBE-07-2020-0111

LangChain. 2022. LangChain [Software]. Retrieved from https://github.com/langchain-ai/langchain

Lee, J.H., Yi, J. & Son, J. 2019. Development of an Automatic-Extraction Model of Poisonous Clauses in International Construction Contracts Using Rule-Based NLP. *Journal of Computing in Civil Engineering*, 33.

Manoliadis, O. G. & Vatalis, K. I. 2016. Sustainability criteria for awarding construction contracts in Greece. *International Journal of Sustainability Management and Information Technology*, 2(2), 7–12.

Mehrbakhsh Nilashi, Rupani, P. F., Rupani, M. M., Kamyab, H., Shao, W., Ahmadi, H., Rashid, T. A. & Aljojo, N. 2019. Measuring sustainability through ecological sustainability and human sustainability: A machine learning approach. *Journal of Cleaner Production*, 240, 118162. https://doi.org/10.1016/j.jclepro.2019.118162

Meta AI. 2025. Llama 4 Multimodal Intelligence [Blog post]. Retrieved from https://ai.meta.com/blog/llama-4-multimodal-intelligence/

Mitkus, S. & Trinkūniene, E. 2008. Reasoned decisions in construction contract evaluation. *Technology in Economy Development*, 14(3), 402–416. https://doi.org/10.3846/1392-8619.2008.14.402-416

Ollama. 2023. Ollama: Run LLMs Locally [Software]. Retrieved from https://ollama.com

Spence, R. & Mulligan, H. 1995. Sustainable development and the construction industry. *Habitat International*, 19(3), 279–292.

Streamlit Inc. 2023. Streamlit [Software]. Retrieved from https://streamlit.io