

BUILDING WITH WORDS: EXPLORING THE OPPORTUNITIES AND CHALLENGES OF USING LARGE LANGUAGE MODELS IN CONSTRUCTION PROJECT MANAGEMENT

Dorđe Nedeljković¹, Nikola Knežević¹, Aleksandra Parezanović¹

¹Faculty of Civil Engineering, University of Belgrade, Serbia

Abstract: The construction industry can benefit significantly from using Large Language Models (LLMs), such as GPT models. LLMs have the potential to significantly improve various aspects of Construction Project Management (CPM), including project scheduling, regulatory compliance, risk management, project monitoring and reporting, contract management, safety protocols, documentation analysis, dispute resolution and budgeting. These models can enhance efficiency by streamlining time-consuming tasks, providing valuable insights from vast textual data, and improving decision-making processes. However, their adoption also poses challenges related to data quality, model biases, ethical considerations, and the need for experts and specialized training for users. The opportunities and challenges of using LLMs in CPM are highlighted, and a use case is illustrated for documentation analysis.

Keywords: GPT models, Artificial Intelligence (AI), Generative AI, Natural Language Processing (NLP), construction industry

1. INTRODUCTION

The construction industry, known for its slower uptake of innovation than other sectors, faces challenges due to its culture and product complexity (Gambatese and Hallowell, 2011). Information plays a vital role in project success, but a lack of integration and efficient management hinders collaboration and productivity. With embracing advancements in technology, the industry is turning to tools like Building Information Modeling (BIM), the application of big data analytics, offsite construction, automation, and artificial intelligence (AI).

AI encompasses a broad range of technologies that enable machines to perform tasks typically requiring human intelligence. Natural Language Processing (NLP) is a crucial component of AI, focusing on enabling machines to understand and generate human language. Large Language Models (LLMs), such as GPT (Generative Pre-trained Transformer), are a subset of AI focused on understanding and generating human-like text based on the input they receive. These models are trained on vast amounts of text data, allowing them to generate coherent, contextually relevant responses (Patil et al., 2024).

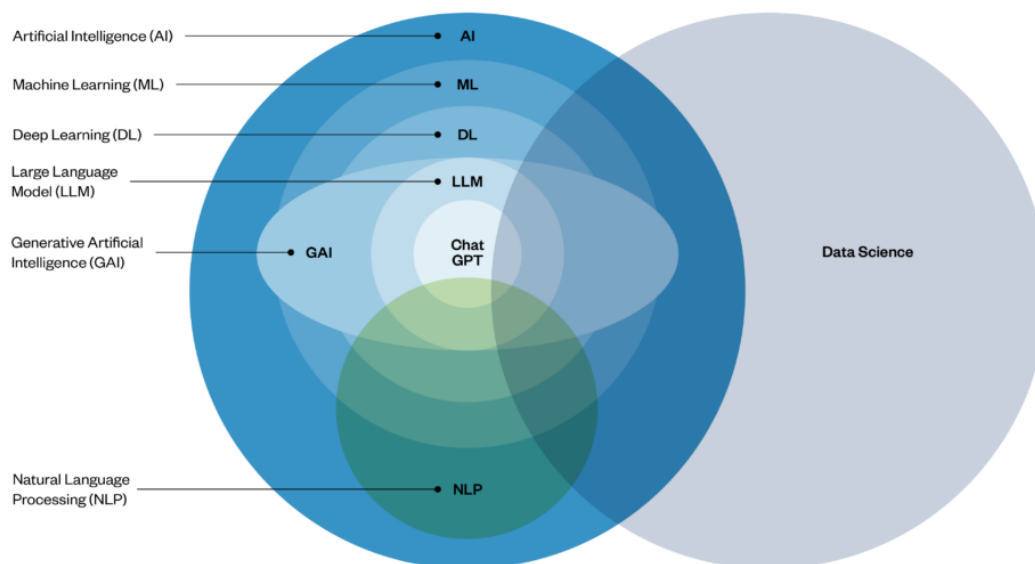


Figure 1. The relationship between LLMs and other types of AI (Knowable, 2024)

LLMs function by analysing the relationships between words in a sentence and across a vast text corpus. They use this understanding to predict and generate text sequences, making them valuable tools for automating communication and data analysis tasks in Construction Project Management (CPM). Their ability to quickly process and synthesize information can help project managers with their daily tasks, including scheduling, risk assessment, resource allocation, and stakeholder communication (Saka et al., 2023)

To enhance LLMs in CPM, the Retrieval-Augmented Generation (RAG) technique is effective. RAG combines generative capabilities with a retrieval system. When a query is inputted, like about contract terms or compliance, the model can be used to retrieve information from relevant documents from company's archives from which it has access to. This ensures information is specific to the company's operations, aiding accurate responses. Implementing RAG involves steps like organizing data, integrating with existing IT infrastructure, and ensuring maintenance. While it can improve efficiency, RAG requires investment and depends on existing data infrastructure and technical expertise.

2. LITERATURE REVIEW

AI has been leveraged in cost prediction, delay prediction, building design, energy prediction, workers' activity recognition, construction site safety, cash flow prediction, structural health monitoring, resource allocation and optimization, predictive maintenance, and decision support systems, among others (Abioye et al., 2021). However, LLMs, such as GPT models, are still new in the construction industry. This chapter presents findings from the literature search on the current applications of GPT models in the CPM.

2.1. CURRENT APPLICATION

Effective CPM relies heavily on efficient scheduling and logistics handling, which can be particularly challenging due to the unpredictable and complex nature of the supply chain. LLMs have demonstrated their ability to acquire and analyze textual information, offering a deeper understanding of the intricacies involved in project scheduling and logistics. Furthermore, LLMs aid in resource allocation and optimization by providing insights into project requirements, available resources, and potential challenges along the critical path. They facilitate scenario analysis for documenting resource allocation strategies (Saka et al., 2023). LLMs also enhance progress monitoring and reporting by analyzing textual data related to projects, offering a comprehensive solution for automation during construction (Shamshiri et al., 2024). Moreover, these models can recognize and analyze various types of documents, establishing a centralized database for easy access by stakeholders, enabling effortless generation of progress reports and status updates from extracted data. In managing changes to project scope, LLMs assist in reviewing change orders, ensuring consistency with contract documentation, and facilitating communication during negotiation and dispute resolution (Abioye et al., 2021). Additionally, LLMs can forecast conflicts or claims in construction projects based on past records, proposing tailored dispute resolution strategies to minimize legal battles (Kalogeraki et al., 2024). In traditional approaches, regulatory compliance verification tends to be time-consuming, error-prone, and resource-intensive. However, LLMs, with their NLP capabilities, can streamline this process by automating compliance checks, thereby reducing time and effort (Beach et al., 2020). In safety management, integrating LLMs enables efficient safety audits, automated risk assessments, and knowledge dissemination through interactive training for workers (Porter, 2021). Risk management is crucial during the construction phase, and while traditional methods have limitations, LLMs show promise in enhancing hazard identification, assessment, and management. By automating the generation of risk reports and communication, these models improve accuracy and consistency, minimizing misunderstandings and omissions (Phasha et al., 2022).

2.1.1. DATA TYPES HANDLED BY LLMs

LLMs possess exceptional proficiency in processing and generating textual content, a capability essential to the majority of their application. In CPM, textual data in this context encompasses various forms such as project correspondence, including emails, memos, and communication with stakeholders; contracts and legal documents like agreements, terms of service, and compliance documents; as well as project reports and documentation such as progress reports, minutes of meetings, and planning documents.

LLMs, while primarily focused on text processing, can also handle structured data such as tables, spreadsheets for budgets, schedules, and resource allocation (Fang et al., 2024). However, when data types exceed LLMs' capabilities, integration with specialized AI systems like Computer Vision (CV) becomes crucial. For instance, CV models can analyze diagrams and photos, aiding LLMs in generating descriptions or extracting insights. Additionally, IoT devices collecting on-site data can be analyzed by LLMs and used for maintenance or safety concerns. Emerging AI technologies, like multimodal learning, further enhance LLM capabilities by integrating text, images, and voice-to-text translation for comprehensive data analysis and understanding.

3. ILLUSTRATIVE CASE

This section explores the practical application of LLMs in CPM. We aim to demonstrate how LLMs can enhance information retrieval and decision-making processes in this field by examining detailed discussions and decisions from actual project meetings. The following interactions, extracted from Weekly Progress Meetings, illustrate the model's capabilities in understanding complex management scenarios, highlighting the opportunities and challenges LLM integration presents in managing large-scale construction projects. This illustrative study serves as a point of reference for assessing the potential of language models to support and transform the construction industry. The study was conducted using ChatGPT-4 Turbo, a state-of-the-art language model developed by OpenAI, due to its top-ranking status during the research period (see Figure 2).

Rank* (UB)	Model	Arena Elo
1	GPT-4-Turbo-2024-04-09	1259
2	GPT-4-1106-preview	1253
2	Claude 3 Opus	1251
2	Gemini 1.5 Pro API-0409-Preview	1250
2	GPT-4-0125-preview	1247
6	Llama-3-70b-Instruct	1210
6	Bard (Gemini Pro)	1209
7	Claude 3 Sonnet	1201
9	GPT-4-0314	1190
9	Command R+	1191
11	Claude 3 Haiku	1181

Figure 2. Global ranking of LLMs by Elo rating system (Huggingface, 2024)

Figure 2 shows the global ranking of LLMs by the Elo rating system, a method used to calculate the relative skill levels of entities in competitor-versus-competitor games. Each LLM starts with an initial rating, and as they “compete” against each other or a set of benchmarks, their ratings are adjusted based on their performance. This provides a dynamic leaderboard where the rankings reflect the relative abilities of the models to understand, generate, or interact with human language effectively.

3.1. TESTING SCENARIOS

Within this subsection, three testing scenarios are explored, wherein questions regarding discussions documented in the MoMs (Qi) are paired with generated answers (Ai) and original content extracted directly from the MoMs documents.

Q1: Did the Employer confirm that the terraces and facade outside the mock-up apartments will not be considered under the subject of approval of mock-up apartments?

A1: Yes, the Employer confirmed that the subject of approval of mock-up apartments should only consider the quality of completed works in the apartments and access hallways, not including the processing of terraces and facades. This was discussed during the Weekly Progress Meeting with reference number MOM_014.

Original Content from MOM_014:

The client confirmed that the subject of approval of mock-up apartments should be the quality of completed works in apartments and access hallways and that the processing of terraces and facades will not be considered.

Q2: When will the Contractor send the claim due to a mitigation plan?

A2: The Contractor stated they would send a claim related to the mitigation plan for compensating the delays caused by the sewage overflow in February 2024. This was discussed during the Weekly Progress Meeting with reference number MOM_029. The Contractor mentioned that this claim would have a financial aspect due to the additional efforts and resources invested to counteract the delays. The assessment of whether the next milestone is achieved will also factor into the formulation of this claim.

Original Content from MOM_029:

The Contractor stated that all delays will be eliminated by the beginning of February 2024, up to the next milestone. The Contractor stated that they reduced the delays by one month. There was a delay of 17 days due to the sewage spill, and they invested additional effort and resources to achieve this. The claim will have a financial aspect due to the mitigation plan, i.e., compensating delay due to the sewage overflow, and it will be sent in February 2024, when it will be seen whether the next milestone is reached.

Q3: Will there be duplicate nominations of PC items? Can the Contractor also offer materials for PC items that are not nominated?

A3: In the meeting minutes with reference number **MOM_029**, it was confirmed that there would be no duplicate nominations of PC (**Permanent Construction**) items suppliers. The Employer has informed us that the Contractor will also be allowed to offer materials for PC items that are not nominated. It was mentioned that in such cases, clear criteria should be defined for selecting equipment and materials. This arrangement allows the Contractor some flexibility in proposing alternative materials or equipment for construction parts, which may help streamline the procurement process and possibly optimize costs or adapt to available resources.

Original Content MOM_016:

The engineer stated that there will be no duplicate nominations of PC items suppliers. Mock-up nomination is final and valid for the whole project unless Employer has a request to change.

Original Content **MOM_024**:

The contractor stated that the deadline for the nomination of PC items has passed. Employer has informed that Contractor will also be allowed to offer materials for PC items that are not nominated. In this case the Contractor requested that clear criteria be defined for selecting equipment and materials.

The answers and referenced sources were correct in the first two testing scenarios. However, in the more complex third scenario, where the correct answer requires combining multiple sources, the model incorrectly referenced sources. Instead of quoting MOM_016 and MOM_024, the model quoted MOM_029, where no relevant content occurred. Additionally, in the third scenario, the model misinterpreted PC items as Permanent Construction items rather than Prime Cost Items, despite the contextual clarity provided in the MoMs. Furthermore, the proposed arrangement in the third scenario has been elaborated upon, despite this explanation not being included initially in the MoMs.

4. DISCUSSION

Implementing LLMs in CPM present both significant opportunities and challenges. On one hand, LLMs offer the potential to significantly improve various aspects of CPM, including regulatory compliance, risk management, progress monitoring and reporting, contract management, documentation analysis and safety management. These models can enhance efficiency by automating daily time-consuming tasks, and facilitating improved decision-making processes. For example, LLMs can streamline progress monitoring and reporting by analyzing textual data from project updates, reports, and documentation, allowing project managers to quickly assess project status and make informed decisions. In terms of contract management, LLMs can assist in reviewing and analyzing contract terms, conditions, and risk identification. Moreover, as shown in the illustrative study, LLMs can play a crucial role in documentation analysis by extracting relevant information from project documents and reports, saving time and effort for project teams. Also, the incorporation of Retrieval-Augmented Generation (RAG) can present additional opportunities by leveraging past project data to improve the information retrieval process, thereby providing more accurate and contextually insights. Finally, in safety management, LLMs can analyze safety reports and identify potential risks on construction sites, contributing to the overall safety culture of the project.

On the other hand, implementing LLMs in CPM faces significant challenges related to data quality. Ensuring accuracy, relevance, and representativeness of the training data is crucial. While LLMs can handle general patterns, they may struggle with domain-specific data. Additionally, LLMs may inherit biases from the training data, leading to hallucinations, where the model generates factually incorrect or nonsensical information. For instance, in the third

testing scenario of the case study, misinterpreting PC items as Permanent Construction items instead of Prime Cost Items, and falsely identifying sources of an answer are the examples of hallucinations. To mitigate hallucinations, it's essential to train the model on high-quality, domain-specific datasets, and implement response verification systems to cross-check outputs against trusted sources. Furthermore, effective prompting plays a crucial role in leveraging the potential of LLMs. Prompting involves inputting text into an LLM to elicit a desired response. To achieve optimal results, it is essential to use clear and specific requests to minimize misunderstandings, utilize follow-up prompts to enhance response refinement for relevance and accuracy, and steer clear of ambiguous language to decrease the chances of generating irrelevant outputs.

Ethical considerations also pose a significant challenge in LLM implementation. While using internally implemented RAG can enhance privacy and data security, as opposed to relying on third-party cloud-based LLMs, there remain inherent risks. For example, even with internal implementations, sensitive project documents could still be exposed if not properly managed. Ensuring transparency, accountability, and fairness in the use of these advanced LLMs, including RAG, is essential to mitigate these ethical risks and safeguard against potential privacy breaches.

The integration of LLMs into existing CPM workflows also requires specialized training for users. Project teams need to develop skills in data organization, model integration, and maintenance to effectively leverage LLMs. Additionally, there may be resistance to change from stakeholders accustomed to traditional project management methods, necessitating careful change management strategies.

Despite the challenges posed by implementing LLMs in CPM, the opportunities they offer are significant. By addressing these challenges through careful planning, and oversight, organizations can harness the full potential of LLMs to improve project outcomes, enhance collaboration, and drive innovation in the construction industry.

5. CONCLUSION

In Construction Project Management (CPM), Large Language Models (LLMs) are now serving effectively as personal assistants, offering substantial support with daily tasks. However, it's vital to consistently verify the accuracy and reliability of the results they provide. The potential applications of LLMs across various sectors within the construction industry are considerable. Therefore, it's crucial to continuously monitor advancements in this area, as integrating LLMs into CPM processes is likely to become increasingly necessary. Ultimately, as the capabilities of LLMs continue to expand, they have the potential to transform the construction industry by enhancing efficiency, reducing errors and time-consuming tasks, and facilitating more informed decision-making.

LITERATURE

1. Abioye, S.O., Oyedele, L.O., Akanbi, L., Ajayi, A., Davila Delgado, J.M., Bilal, M., Akinade, O.O., Ahmed, A. (2021). Artificial intelligence in the construction industry: A review of present status, opportunities and future challenges. *J. Build. Eng.* 44 <https://doi.org/10.1016/j.jobbe.2021.103299>.
2. Beach, T. H., Hippolyte, J. L., & Rezgui, Y. (2020). Towards the adoption of automated regulatory compliance checking in the built environment. *Automation in construction*, 118, 103285.
3. Ezelogs, 2023. Ultimate & Exciting Decision Making Using Chat GPT in Construction. <https://blogs.ezelogs.com/decision-making-using-chat-gpt/>.
4. Fang, X., Xu, W., Tan, F. A., Zhang, J., Hu, Z., Qi, Y., ... & Faloutsos, C. (2024). Large Language Models on Tabular Data--A Survey. *arXiv preprint arXiv:2402.17944*.
5. Gambatese, J. A., & Hallowell, M. (2011). Factors that influence the development and diffusion of technical innovations in the construction industry. *Construction Management and Economics*, 29(5), 507–517. <https://doi.org/10.1080/01446193.2011.570355>.
6. Huggingface , 2024. “LMSYS Chatbot Arena Leaderboard”, Retrieved May 4, 2024, from <https://huggingface.co/spaces/lmsys/chatbot-arena-leaderboard>.
7. Kalogeraki, M., & Antoniou, F. (2024). Claim Management and Dispute Resolution in the Construction Industry: Current Research Trends Using Novel Technologies. *Buildings*, 14(4), 967. <https://doi.org/10.3390/buildings14040967>.
8. Knowable (2024). “Introduction: How to Apply LLMs and AI to Contracts”. <https://www.knowable.com/blog/introduction-how-to-apply-llms-and-ai-to-contracts/>.
9. Patil, R., & Gudivada, V. (2024). A Review of Current Trends, Techniques, and Challenges in Large Language Models (LLMs). *Applied Sciences*, 14(5), 2074.
10. Phasha, C. (2022). The impact of Artificial Intelligence on cost overruns and risk management in Construction Project Management (Doctoral dissertation, University of Johannesburg).
11. Porter, S., (2021). Construction Industry Looking to Predictive Analytics to Improve Project Outcomes. <https://www.oracle.com/news/announcement/idc-construction-industry-looks-to-predictive-analytics-to-improve-project-outcomes-2021-10-27/>.
12. Saka, A., Taiwo, R., Saka, N., Salami, B. A., Ajayi, S., Akande, K., & Kazemi, H. (2023). GPT models in construction industry: Opportunities, limitations, and a use case validation. *Developments in the Built Environment*, 100300.
13. Shamshiri, A., Ryu, K. R., & Park, J. Y. (2024). Text mining and natural language processing in construction. *Automation in Construction*, 158, 105200.
14. Uddin, S.J., Albert, A., Ovid, A., Alsharef, A., (2023). Leveraging ChatGPT to Aid construction hazard recognition and support safety education and training. *Sustainability* 15 (9), 7121.