

Article

Optimizing the Utilization of Generative Artificial Intelligence (AI) in the AEC Industry: ChatGPT Prompt Engineering and Design

Reihaneh Samsami

Department of Civil and Environmental Engineering, College of Engineering, University of New Haven, West Haven, CT 06516, USA; rsamsami@newhaven.edu

Abstract: Generative Artificial Intelligence (AI) holds significant potential for revolutionizing the Architecture, Engineering, and Construction (AEC) industry by automating complex tasks such as construction scheduling, hazard recognition, resource leveling, information retrieval from BIM, etc. However, realizing this potential requires a strategic approach to ensure effective utilization and maximum benefit. This paper presents guidelines for prompt design and engineering to elicit desired responses from ChatGPT, a Generative AI tool, in AEC applications. Key steps include understanding user intent, leveraging model capabilities, and optimizing prompt structures. By following these guidelines, stakeholders in the AEC industry can harness the power of Generative AI to improve construction scheduling processes, increase project efficiency, and ultimately drive innovation and growth in the industry. Several illustrative examples on construction scheduling and hazard recognition are provided to demonstrate the methodology proposed in this research. It is concluded that Generative AI, when effectively utilized, significantly enhances project scheduling and hazard recognition capability in the AEC industry with minimal error.

Keywords: generative AI; ChatGPT; prompt design and engineering; construction scheduling; hazard recognition; AEC



Citation: Samsami, R. Optimizing the Utilization of Generative Artificial Intelligence (AI) in the AEC Industry: ChatGPT Prompt Engineering and Design. *CivilEng* **2024**, *5*, 971–1010. <https://doi.org/10.3390/civileng5040049>

Academic Editors: Ying Huang and Angelo Luongo

Received: 12 June 2024

Revised: 23 September 2024

Accepted: 17 October 2024

Published: 28 October 2024



Copyright: © 2024 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In the Architecture, Engineering, and Construction (AEC) industry, the integration of Generative Artificial Intelligence (AI) tools such as ChatGPT holds transformative potential. These advanced technologies promise substantial improvements in areas like construction scheduling and hazard recognition. However, their practical application is currently restricted due to generalized interaction methods and insufficient prompt engineering. This frequently results in outputs that are not adequately aligned with the specific requirements of the AEC sector, limiting both the effectiveness and acceptance of generative AI technologies within the industry.

Effective utilization of Generative AI in the AEC field requires specialized prompt engineering—the careful crafting of queries that guide AI to generate precise and relevant outputs [1]. Mastering this art of prompt engineering is essential to surmount the limitations of standard AI responses, thereby aligning AI capabilities with the distinct tasks of the AEC industry to enhance the relevance and accuracy of the information provided. While ChatGPT has demonstrated strong performance on general benchmark datasets, there is a distinct lack of empirical validation for its use in construction-specific tasks, such as scheduling and hazard recognition.

The objective of this paper is to develop structured guidelines for prompt engineering that are specifically tailored to meet the needs of the AEC industry. By designing and implementing well-considered prompts, this research aims to optimize the responses from ChatGPT for practical applications, thus improving decision-making processes. Importantly, the guidelines are crafted to be accessible even to those in the AEC industry with

limited experience with ChatGPT, ensuring broad usability. The proposed methodology is evaluated through detailed case studies in two domains: construction scheduling and hazard recognition. Ultimately, the performance of ChatGPT is measured using an accuracy metric named “Relative Error”.

2. Literature Review

This section critically examines recent studies on the application of Generative AI in the AEC industry, focusing on how these technologies have been adapted to address specific challenges in the AEC domain.

2.1. Natural Language Processing (NLP)

To complete a successful construction project, stakeholders must consider the experience from previously completed projects. As part of this, the first step is to collect and store all sorts of data related to the AEC industry [2]. Further, this data can be fine-tuned and used for any future projects. Fine-tuning is the process by which the system generates appropriate responses for specific prompts. For this task, the system must process certain tasks to understand the data. Natural Language Processing (NLP), a branch of AI, can understand and process text and deliver it in a human-understandable way. In real time, NLP can process text from one language to another and can respond to and command a large volume of text instantly. In the AEC industry, NLP is a tool that processes and extracts auto-generated text from several generative AI platforms like ChatGPT, Gemini (formerly Bard), and Copilot. Generative AI in the AEC industry works with the formal approach of NLP.

OpenAI’s Generative Pre-Trained Transformer (GPT): ChatGPT

Generative Pre-trained (GPT) models from OpenAI have advanced the field of NLP. An OpenAI language model, Generative Pre-Trained Transformer (GPT), can produce responses to text that are almost identical to natural human language [3]. The first model, GPT-1, was released in June 2018, and the latest edition, GPT-4, was released in March 2023 (Figure 1). Refining the concepts behind GPT models involves two steps: discriminative supervised fine-tuning to enhance performance on particular tasks, and generative unsupervised pretraining using unlabeled data [4]. Machine learning algorithms are now capable of training from massive amounts of data and identifying complex structures in simple language without the need for programming because of the development of data-driven techniques like GPT models [5]. The model develops organically during the pretraining phase, much like a human would do in a novel setting, but the creators train it more formally during the fine-tuning phase [6].

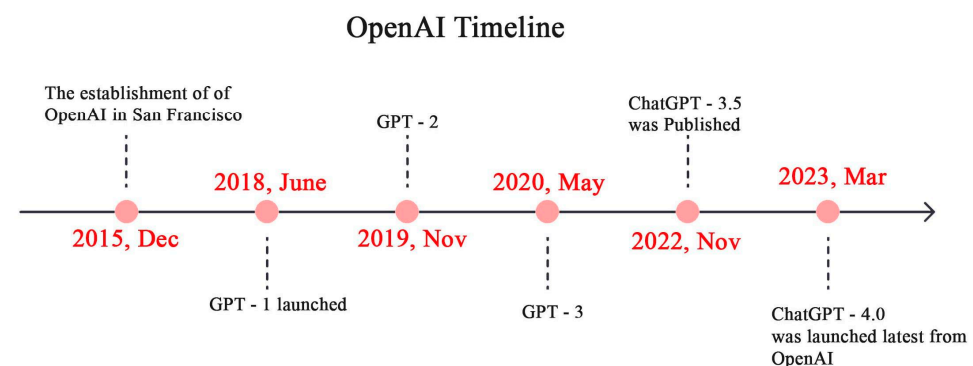


Figure 1. Progress of GPT models.

2.2. Potential Applications of Generative AI in AEC

While Generative AI is being utilized in some industries like healthcare, its application is very limited in the AEC industry. This might be due to a setback in large data sets and

the nature of AEC industry-specific prompts [7]. To explore the potential applications of Generative AI, the authors have reviewed the few available studies on this topic. These studies have utilized Generative AI at different phases of the project lifecycle, including planning [8], material selection [5], workforce safety education [9], Building Information Modeling (BIM) [7], scheduling [10], and sequence planning [8]. These applications are briefly summarized in the following subsections.

2.2.1. Information Retrieval from BIM

BIM is a widely used tool in the AEC industry and represents one of the most significant digital assets throughout a project's lifecycle. BIM supports construction and facility management operations by integrating data from various disciplines, including architectural, structural, mechanical, electrical, and plumbing. BIM's cloud-based data and information specific to the AEC industry can be used to train GPT models, potentially reducing time and costs while increasing the efficiency of maintenance and operations. Thus, BIM is considered an optimal solution for information retrieval and interaction with GPT models [7].

In this context, Zheng et al. [7] proposed a methodology (Figure 2) to combine BIM with GPT to achieve efficient results through their developed prompt-based Virtual Assist (VA) framework. As part of the framework, the authors developed a prompt library and a prompt manager, enabling GPT to respond to user queries related to BIM. Additionally, the BIM-GPT framework is designed to provide a Natural Language (NL) interface equipped with BIM visualizations, facilitating effective information retrieval and interaction with 3D building components.

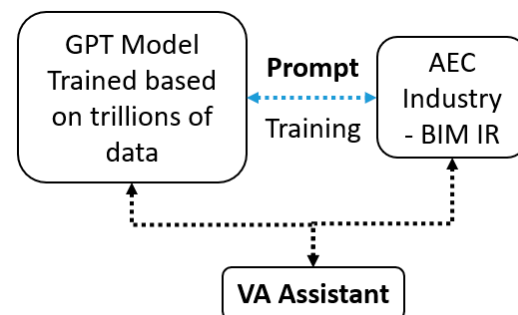


Figure 2. Information retrieval from BIM Methodology, recreated from [7].

2.2.2. Robotic Sequence Planning

The application of robotic technologies and accompanying modifications to the current construction process can greatly increase productivity, lower construction costs, and enhance project safety [9]. Robot-based construction offers several potential advantages, but one of its biggest challenges is the need for precise and effective sequence planning. Consequently, construction personnel often have to perform manual sequence planning on the job, determining the best sequence for building phases and addressing related logistics issues [8].

A study by You et al. [8] presents RoboGPT, a new approach (Figure 3) for automated sequence planning in robot-based assembly applicable to construction activities, leveraging ChatGPT. As one of the most advanced language models, ChatGPT has significant potential to understand and utilize Large Language Models (LLMs) to generate instant, logical, human-like text for sequence planning in the construction industry. Based on the sequence planning generated by ChatGPT, instructions were provided to the robot to execute tasks in the specified order.

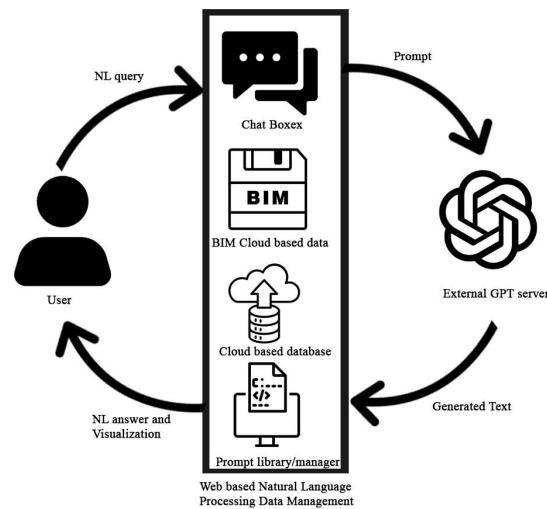


Figure 3. Robotic sequence planning methodology, recreated from [8].

2.2.3. Scheduling

The introduction of GPT into the construction industry has the potential to automate labor-intensive and time-consuming tasks. Prieto et al. [10] conducted research on using GPT to generate construction schedules based on NLP. To evaluate the results, the authors used several parameters, including efficiency, clarity, coherence, reliability, relevance, accuracy, and adaptability. Each parameter was measured by comparing the results generated by GPT with the documents prepared by the project manager.

2.2.4. Hazard Recognition

Uddin et al. [9] investigated the use of ChatGPT to improve hazard recognition in construction and support safety education. Their study addresses the widespread issue of poor hazard recognition in construction workplaces, which can lead to serious safety incidents. The authors conducted an experimental study with 42 students in a construction program, measuring their hazard recognition abilities before and after introducing ChatGPT. The findings indicate that ChatGPT significantly enhances hazard recognition, with students recognizing more hazards post-intervention. The research suggests integrating ChatGPT into safety training and education to better prepare future construction professionals and improve workplace safety. Figure 4 summarizes this study.

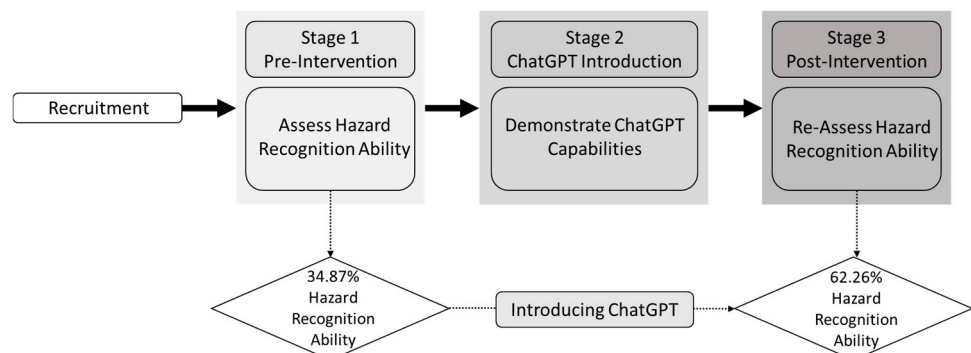


Figure 4. Hazard recognition methodology, recreated from [9].

2.2.5. Master Schedules

Planning tasks and activities in the master schedule are developed and structured differently to meet distinct needs. Typically, planning master schedules are created in spreadsheets with basic work breakdown structure information. Alternatively, a master scheduling software such as Microsoft Project and Oracle P6 is also used. Generally,

project managers create the main activities, while superintendents develop the smaller tasks. Amer et al. [11] conducted a study using multiple approaches for problem-solving. Initially, the authors used a distance-based baseline model, followed by the extraction of semantic location data. Ultimately, they introduced a Machine Learning model built on the state-of-the-art GPT-2.

2.2.6. Material Selection and Optimization

To achieve the practical implementation of GPT in the AEC industry, a research study by Saka et al. [5] presented a Material Selection and Optimization prototype model aimed at testing the efficiency of GPT in the construction sector during the design phase. The study used a BIM model containing several elements as a case. Three key steps were involved in the development process: (a) the data retrieval model, (b) the user interface, and (c) the NLP module [12]. In the data retrieval model, the BIM model was processed to extract 2D and 3D geometry and usable NLP modules. The prompt development using ChatGPT in the NLP module designed an interactive dialogue system through a multi-step process [13]. Finally, with the help of JavaScript, HTML, CSS3, and Bootstrap 5 technologies, Forge Viewer was developed to display the visualizations on the front-end [5].

2.2.7. Other Potential Applications

The other potential applications of ChatGPT in the AEC industry are as follows:

- **Pre-design:** In the pre-design stage, ChatGPT can assist with several aspects such as costing, project scope, timeline, and value engineering. A few studies [5,14–16] explored these capabilities. ChatGPT can provide preliminary cost estimates, define project scopes, and optimize timelines to ensure projects are completed on time and within budget.
- **Design:** During the design phase, ChatGPT can aid in generating design concepts, developing design specifications, ensuring regulatory compliance, performing quantity takeoffs and costing, optimizing material selection, and conducting energy efficiency analyses [17]. For tasks like design concepts and specifications, zero-shot and few-shot approaches have been adopted by [5]. ChatGPT can help navigate the complexities of frequently updated building codes and safety standards, reducing the risk of errors that could affect project schedules and costs. It can also support the preparation of bills of quantities and reference standards and techniques to improve building energy efficiency, such as passive solar design and renewable energy integration.
- **Construction:** In the construction phase, ChatGPT can manage and interpret both numerical and textual data [18]. It can optimize logistics and project scheduling, ensuring that projects are completed on time and within budget [5].
- **Operation and maintenance:** ChatGPT can be leveraged to predefine the lifetime and schedule timely maintenance of machinery and construction equipment, thereby avoiding hazards on job sites [19]. In the event of an incident, ChatGPT can provide instant solutions to manage the situation effectively [5].
- **Demolition phase:** During demolition, ChatGPT can offer safety instructions to crew members to prevent hazardous situations [20]. Additionally, GPT models can assist in waste management, risk assessment, structural analysis, material recovery, and redevelopment planning [5].
- **Value-Added Services:** Although not directly tied to a specific phase of the construction project lifecycle, ChatGPT can enhance various value-added services. These include customer service and support, knowledge management and training, chat-bot/conversational systems, stakeholder management, and market analysis [5].

2.3. Challenges and Limitations of Generative AI

While the use of ChatGPT in the AEC industry offers numerous benefits, there are several challenges and limitations to its implementation, as noted by several studies [5,14–16,18–20] as follows:

- **Hallucination:** Models can provide incorrect information that sounds convincing but is entirely false, reducing performance.
- **Data and Interoperability:** The lack of structured data and the slow transition to digital technology in the industry make it difficult to fine-tune models. Data in various formats (PDF, HTML, CAD, IFC) must be standardized, which can dilute the information.
- **Domain-Specific Knowledge and Regulatory Compliance:** Models have limited understanding of domain-specific knowledge. The construction industry's extensive and varying regulations require comprehensive knowledge and logical reasoning, which models may lack.
- **Confidentiality and Intellectual Property:** Projects contain sensitive data (designs, costs, contracts). There are concerns about maintaining confidentiality and avoiding intellectual property infringements when using AI.
- **Trust and Acceptability:** Resistance to change and skepticism about AI's reliability are prevalent. Industry professionals fear AI may replace jobs and are hesitant to trust AI for significant decisions.
- **Liability and Ethics:** Bias, incomplete information, and inaccuracies can lead to harmful outputs. AI models may struggle to fully understand legal requirements, necessitating clear guidelines and human oversight.
- **Skills and Training:** Effective use of AI requires new skills and training, particularly in prompt engineering. This requirement may slow progress.
- **Infrastructure Requirements and Costs:** Smaller businesses may lack the resources for the necessary computing power, network connectivity, and data storage.
- **Scalability and Performance Optimization:** Fine-tuning models for diverse tasks remains a challenge. Models may underperform outside their trained tasks.
- **Cybersecurity:** Increased digital interconnectedness raises the risk of unauthorized access, data breaches, and cyber-attacks. Proper training and awareness are essential.
- **Interdisciplinary Collaboration:** Effective collaboration across architecture, engineering, and construction management is crucial. New technologies may face skepticism, affecting collaboration.
- **Cultural and Social Considerations:** Fear of job displacement and data privacy concerns can hinder AI acceptance. Social and cultural factors significantly influence trust and adoption.
- **Latency Issues:** Time-sensitive projects may suffer from delays due to data processing and training lag.
- **Maintenance and Upkeep:** AI models require continuous maintenance and updates to prevent degradation and ensure accuracy.
- **Multilingual Language Processing:** There is a growing need for multilingual capabilities. ChatGPT has shown underperformance in languages other than English.
- **Standards and Variability:** Diverse and evolving industry standards across regions and projects complicate AI implementation. Models need adaptive databases updated with the latest standards and codes.

2.4. Prompt Engineering and Design

To achieve effective results from OpenAI models like ChatGPT, it is essential to communicate effectively with them. For instance, to obtain accurate and coherent results from ChatGPT, it is important to provide relevant prompts. This helps to keep the results within the bounds of the relevant datasets. Prompt engineering can be used for this purpose, producing prompts and instructions for OpenAI models.

Prompt engineering is the art of designing relevant inputs for OpenAI models like ChatGPT [1]. The primary role of prompt engineering is to serve as a medium of communication between the end user and ChatGPT [21]. It involves creating a set of instructions in the form of prompts to guide ChatGPT in generating specific outputs [22]. In the AEC industry, the following factors can influence prompt selection:

- **User Intention:** Users anticipated the final output. This helps to meet the end-user requirements by understanding the purpose of interaction with ChatGPT.
- **Model Understanding:** By understanding the strengths and limitations of ChatGPT, this will assist in the design of prompts to provide relevant responses to most of the model's capacities. It should also be noted that it may also produce incorrect responses at times.
- **Domain Specificity:** when the end user is dealing with a specific domain of research like the AEC industry, the end-user should provide relevant prompts to guide the model to generate relevant responses.
- **Clarity:** The end user must ensure that the given prompts are clear enough to avoid any uncertainty that may lead to less ideal responses. This uncertainty may arise from a lack of clarity in the prompts.
- **Constraints:** To achieve the desired output, it is also necessary to determine constraints like the length of response and the required format, etc. Mentioning those constraints will help the model to fine-tune the response and can also meet the expectations of the end user.

These factors (Figure 5) can help design prompts to effectively communicate with ChatGPT and generate quality responses [21]. Improving prompt engineering requires a deeper understanding of the technology's underlying principles, practical experience with systems using the technology, and continual improvement and refinement of skills based on feedback.

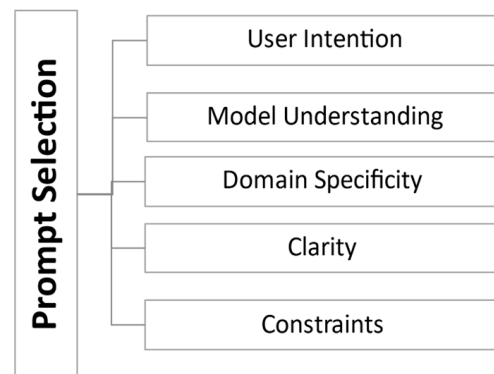


Figure 5. Prompt engineering factors.

The following are tangible measures that professionals can take to strengthen their abilities in prompt engineering:

- **Understanding Fundamental Concepts:** Gaining a basic understanding of how artificial intelligence and machine learning models operate can serve as a foundation for developing prompt engineering skills. This comprehension can be achieved without prior technical or coding skills.
- **Familiarizing with Large Language Models (LLMs):** Understanding the unique strengths and limitations of different LLMs can help generate more effective prompts.
- **Practicing Regularly:** Frequent interaction with LLMs and recording the prompts that yield the most precise and informative responses can be beneficial. As practice makes perfect, this helps in refining prompt engineering skills.
- **Additionally, evaluating prompts in real-world situations is essential to assess their efficiency and effectiveness.**

2.5. Existing Gaps

While it is commonly recognized that prompt engineering with ChatGPT has exhibited exceptional performance in contrast to other leading methods on benchmark datasets [23], there exists a notable absence of empirical research and validation studies concerning construction-related tasks, such as scheduling and hazard recognition, utilizing ChatGPT.

Despite the broad spectrum of potential applications in the AEC domain, limited studies have emphasized the importance of assessing the applicability and validity of ChatGPT and prompt engineering within this context.

Furthermore, the few existing studies tend to focus on isolated tasks without exploring the integration of ChatGPT into larger, more complex workflows commonly found in the AEC industry. As construction projects often involve a dynamic interplay of multiple factors, such as scheduling, resource management, and safety regulations, it becomes crucial to validate AI outputs in the context of such multifaceted environments. The lack of research on ChatGPT's performance in addressing these interconnected tasks creates uncertainty regarding its practical effectiveness.

Additionally, current studies rarely address the need for continuous fine-tuning and adaptation of generative AI models like ChatGPT to accommodate evolving construction standards and project-specific requirements. This gap highlights the need for future work that not only evaluates AI's immediate performance but also examines how these systems can be adapted over time to remain relevant in a constantly changing industry.

Given the critical role of safety and compliance in construction, a significant gap also exists in evaluating ChatGPT's ability to reliably recognize hazards and offer safety recommendations that align with the latest industry standards. As regulations and safety protocols continue to evolve, the ability of generative AI to stay up to date without human oversight remains a key concern that requires further investigation.

3. Materials and Methods

Given the research objective and existing gaps, a systematic approach is designed (Figure 6) to optimize the use of ChatGPT4 through prompt engineering. The process begins with a comprehensive literature review to identify effective prompt engineering techniques across various disciplines. This review aims to select strategies that enhance ChatGPT4's responses by improving prompt clarity, context, and specificity. The methodology then progresses to designing and testing prompts through controlled experiments, comparing ChatGPT outputs against case study references using precise evaluation metrics. Finally, the process concludes with analyzing the effectiveness of these techniques and refining prompts to optimize ChatGPT4's response accuracy and relevance.

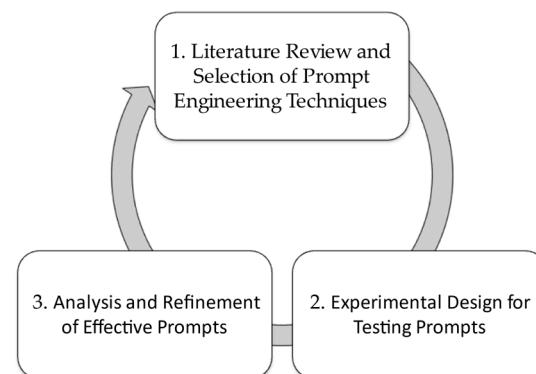


Figure 6. Proposed prompt engineering methodology.

3.1. Literature Review and Selection of Prompt Engineering Techniques

The initial phase of the methodology involves a literature review aimed at collecting existing knowledge on prompt engineering techniques from a variety of disciplines. This review seeks to identify strategies that have proven effective in eliciting enhanced responses from ChatGPT. Emphasis is placed on techniques that improve the clarity, context, and specificity of prompts. Subsequent to this review, a selection of techniques believed to be most promising for augmenting the performance of ChatGPT in complex task scenarios within the AEC industry will be made.

3.1.1. Zero-Shot vs. Few-Shot Prompts

The first concept to understand is the difference between Zero-Shot and Few-Shot prompts. In Zero-Shot prompts, the prompt itself is not providing ChatGPT with instruction or further information to guide the model towards a desirable output. While, on the other hand, Few-Shot prompts include additional information for the model to lead it to the desired direction [24].

3.1.2. Prompt Optimization Techniques

There are several techniques in the literature for optimization of prompts. The authors have selected the following techniques for implementation in this research and a scheduling-related example is given for each technique [25]:

- **Instructional Prompts:** Incorporating instructions on format, style, etc. to the prompt to guide ChatGPT towards the desired output.
Example: "Write a 500-word essay on the importance of effective scheduling in construction projects, focusing on its impact on budget management and deadline adherence. Use a formal tone and include at least two real-world examples to illustrate your points".
- **Constraint-based Prompts:** Defining explicit constraints such as word count limits, required keywords, etc. to control ChatGPT's response and generate output that satisfies specified requirements.
Example: "Provide a brief overview of construction scheduling in exactly 100 words, ensuring you include the keywords 'Gantt chart', 'critical path method', and 'resource allocation'. Make it suitable for an audience with no prior knowledge of construction management".
- **Exemplar-based Prompts:** Creating examples for the model, with specific instances of the desired output, to encourage the model towards generating output that is aligned with the given example, in terms of style, content, structure, etc.
Example: "Here's an example of a concise explanation: 'Construction scheduling is the process of mapping out the start and completion dates of various tasks in a construction project to ensure timely completion. Tools like Gantt charts help visualize this timeline'. Now, write a similar brief explanation focusing on the role of technology in enhancing construction scheduling, maintaining the concise and informative style".
- **Contextual Prompts:** Including relevant information in the prompt to generate coherent output in the context provided.
Example: "Considering a scenario where a construction project is set in a high-traffic urban area with limited working hours due to noise regulations, describe how scheduling adjustments can be made to minimize disruptions and still meet the project deadline".
- **Priming Prompts:** Seeding the model with the primary portion of the desired output, to generate more contextually appropriate outcomes.
Example: "Effective construction scheduling involves meticulous planning and flexibility. Given this statement, expand on how incorporating buffer times for unforeseen delays and regular progress reviews can contribute to maintaining schedules despite unexpected challenges".
- **Reformulating and Rewriting Prompts:** Modifying the structure of a prompt to guide the model in the desired direction and create more accurate output.
Example: Change the Original Prompt: "Tell me about construction scheduling" to Revised Prompt: "Explain the process and significance of scheduling in construction projects, highlighting how it influences project efficiency, cost control, and stakeholder satisfaction".
- **Setting Prompt Variations:** Different variations such as role, tone, and temperature can be defined to improve the quality of the generated output.
Example: "Write a summary of construction scheduling techniques from the perspective of a seasoned project manager emphasizing practicality and problem-solving.

Then, rewrite the summary from the perspective of a new construction management graduate, focusing on theoretical understanding and learning enthusiasm”.

3.1.3. Bad Prompts

There are also some bad prompts that lead to undesired output. Some of these bad prompts [24] are:

- **Vague and Misleading Prompts:** ChatGPT is not able to generate a meaningful output for vague prompts. It also lacks the ability to generate fully unbiased responses, when a question is leading towards a specific outcome.
Example: Change “Do something about scheduling” to “Explain the steps involved in creating a construction project schedule, including key considerations for task sequencing and resource allocation”.
- **Mathematical and Logical Prompts:** It is recommended to ask ChatGPT to “think step by step” when dealing with mathematical or logical questions.
Example: Change “How quickly can a building be constructed?” to “Considering a 10-story office building with a total floor area of 100,000 square feet, estimate the construction schedule assuming standard working hours, a crew of 50 workers, and typical urban construction conditions. Break down the process step by step, including key phases like design, permitting, foundation work, structural framing, and finishing”.
- **Academic References:** It is not recommended to ask ChatGPT for references, especially academic citations, as the model is prone to generate fake output.
Example: “Provide a detailed analysis of the most effective scheduling techniques according to Jones & Smith (2023)”.

3.1.4. OpenAI Prompt Engineering Strategies

OpenAI [26] has collected a series of strategies for prompt engineering. These strategies, including tactics and a scheduling-related example for each tactic, are listed as follows:

Strategy 1—Precise Instruction Writing: Crafting prompts with explicit detail regarding desired response length, complexity, and style ensures responses are closely aligned with expectations.

- Tactic 1.1: Enrich your query with specifics to elicit more relevant responses.
Example: “How do I efficiently allocate resources while scheduling overlapping phases in a Gantt chart?”
- Tactic 1.2: Encourage the model to assume a specific role.
Example: “Compose a message as a project manager, expressing gratitude to a subcontractor for precise schedule adherence, essential for timeline preservation”.
- Tactic 1.3: Employ clear markers to distinguish different segments of your input.
Example: “Examine the project schedule in the supplied CSV. Pinpoint and propose adjustments for any subcontractor schedule overlaps”.
- Tactic 1.4: Clearly outline the task steps.
Example: “Create a detailed instruction manual for developing a Gantt chart in Excel for a compact renovation, including task dependencies and resource assignment”.
- Tactic 1.5: Illustrate with examples.
Example: “Detail instances of how unexpected delays were navigated in construction projects, outlining stakeholder communication and schedule adjustment strategies”.
- Tactic 1.6: Define the output’s desired extent.
Example: “Condense the essential actions for crafting a construction project timeline into three bullet points, emphasizing planning, resource distribution, and oversight”.

Strategy 2—Utilizing Reference Materials: Supplementing queries with reference texts minimizes inaccuracies, enabling models to produce responses with greater accuracy and context.

- Tactic 2.1: Direct the model to base answers on provided texts.
Example: “Determine best practices for handling supply chain issues from the construction guide in the PDF”.
- Tactic 2.2: Apply embeddings for effective information retrieval.
Example: “Search the integrated construction regulation database for high-rise safety standards”.
- Tactic 2.3: Request answers that reference the supplied text.
Example: “Using the urban development code document, describe the process for securing a permit for a new commercial structure, citing each phase”.

Strategy 3—Breaking Down Complex Tasks: Segmenting complicated tasks into simpler, more manageable components increases precision and effectiveness through stepwise engagement.

- Tactic 3.1: Classify queries to apply the most pertinent instructions.
Example: “Identify the construction inquiry as concerning ‘material logistics’, ‘crew scheduling’, or ‘compliance with safety’, providing targeted advice accordingly”.
- Tactic 3.2: Summarize or distill extended dialogues.
Example: “Condense the main points from a detailed discussion on project oversight, focusing on consensus action points concerning schedule management”.
- Tactic 3.3: Sequentially summarize extensive documents.
Example: “Abstract each section of the project management guide, then synthesize a comprehensive summary capturing essential management insights”.

Strategy 4—Allowing Time for Thought: Encouraging a process of stepwise contemplation aids in generating more precise and considerate responses.

- Tactic 4.1: Prompt the model to devise its solution before conclusion.
Example: “Forecast the completion timeline for the construction phases, taking into account variables like team size and material supply, and then assess against the original schedule”.
- Tactic 4.2: Mask the reasoning process with targeted inquiries or internal dialogue.
Example: “Ascertain the optimal deployment of construction machinery for the week ahead, factoring in immediate task requirements, while keeping the decision-making rationale private”.

Strategy 5—Implementing External Instruments: Augmenting tasks with external resources for complex computations or data acquisition enhances the quality of outputs.

- Tactic 5.1: Leverage knowledge search through embeddings.
Example: “Embed up-to-the-minute meteorological data into the planning tool to adaptively modify task schedules according to weather predictions”.
- Tactic 5.2: Conduct precise calculations or interface with APIs through code execution.
Example: “Run a Python program to deduce the best concrete blend, using live data on weather conditions accessed via an external API”.

Strategy 6—Systematic Modification Testing: A methodical evaluation of prompt adjustments ensures consistent enhancement across various scenarios.

- Tactic 6.1: Assess outputs against definitive answers.
Example: “Evaluate the devised construction timetable with established benchmarks for comparable ventures, pinpointing discrepancies and potential for refinement”.

3.2. Experimental Design for Testing Prompts

This phase consists of constructing a series of experiments designed to test the influence of different prompt engineering strategies on the performance of ChatGPT. A controlled experimental setup is established to test the effectiveness of the designed prompts. This includes both control (basic prompts without optimization) and experimental (optimized prompts based on case studies) groups to isolate the effect of prompt engineering on ChatGPT’s performance. The outputs from ChatGPT are compared against the reference

data from the case studies using predefined evaluation metric named “Relative Error”. Relative Error measures the factual correctness of ChatGPT4’s output compared to the reference data. For two applications of construction scheduling and hazard recognition, this metric is defined as follows:

- **Construction Scheduling:** The absolute difference between the estimated project duration and the actual project duration, divided by the actual project duration. used for scheduling cases.

$$\text{Relative Error} = \frac{|\text{Project Duration Per ChatGPT's Calculation} - \text{Project Duration Per Actual Reference}|}{\text{Project Duration Per Actual Reference}} \quad (1)$$

While a smaller relative error indicates that the estimated project duration is close to the actual duration, signifying higher accuracy, a larger relative error indicates that the estimated duration deviates significantly from the actual duration, indicating lower accuracy.

- **Construction Hazard Recognition:** The number of identified hazards by ChatGPT is compared to the actual hazards identified by OSHA and the relative error is calculated as:

$$\text{Relative Error} = \frac{|\text{No. Identified Hazards Per ChatGPT} - \text{No. of Actual Hazards Per OSHA}|}{\text{No. of Actual Hazards Per OSHA}} \quad (2)$$

While a smaller relative error indicates that the number of estimated hazards is close to the actual hazards recognized by OSHA, signifying higher accuracy, a larger relative error indicates that the estimated hazards deviate significantly from the actual hazard recognized by OSHA, thus indicating lower accuracy.

3.3. Analysis and Refinement of Effective Prompts

Following the experimental phase, an analysis is conducted to determine the effectiveness of the prompt engineering techniques in generating outputs that align with the case study references. This involves a detailed comparison of the ChatGPT4 responses with the expected outcomes derived from the case studies. Based on this analysis, the most effective prompts are identified and further refined to enhance their performance. The refinement process will incorporate feedback loops to iteratively improve prompt design, aiming for optimal clarity, context, and specificity. This analysis is critical for determining which prompt modifications yield the most effective solutions to the presented scheduling problems.

4. Application 1: Construction Scheduling

This section applies the designed prompt engineering methodology on five distinct construction scheduling case studies retrieved from the literature. Each case involves utilizing ChatGPT4 to generate construction schedule-related tasks, which is then compared to the actual as-planned schedule documented for the project. The aim is to evaluate the capability of ChatGPT to create viable construction schedules.

- Case Study 1: Residential Building [27].
- Case Study 2: Multipurpose Building [28].
- Case Study 3: Office Building Reconfiguration [29].
- Case Study 4: Business Park Development [30].
- Case Study 5: Luxury Villa [31].

After several iterations using the methodology developed in this research, the following prompts (Table 1) are designed to generate the desired output (Duration, Gantt Chart, and Probabilistic Duration) based on the project information provided to ChatGPT4:

Table 1. Scheduling prompts.

No.	Description	Prompt
PS1	Prompt Designed for Estimating the Duration for Each Activity and the Total Project Duration	<p><i>"I'm working on a "X" construction project and need a concise project schedule, excluding the "Y" phase.</i></p> <ol style="list-style-type: none"> <i>Could you provide a table detailing each phase from the beginning to completion?</i> <i>Include columns for ID, Outline Number, Name, Duration, Predecessors, Start, and Finish dates, starting on day "Z".</i> <i>Also, calculate the project duration in days".</i>
PS2	Prompt Designed for Creating a Gantt Chart for the Project	<p><i>"Please create a Gantt chart for this project, detailing each phase of the project, shown in the table. Ensure the following specifics are incorporated for optimal clarity and presentation:</i></p> <ol style="list-style-type: none"> <i>Column Titles (Dates) Orientation: Incline the date labels to enhance readability.</i> <i>Spacing Between Dates: Adjust the chart to a wider format to prevent date columns from overlapping, ensuring each date fits within its column.</i> <i>Date Simplification: Remove every other date to declutter the view, aiming for a cleaner presentation.</i> <i>Title Update: Change the chart title to 'Gantt Chart for Case Study 1'.</i> <i>Document Format: Include the final Gantt chart in an Excel document (.xls format), ensuring all parts of the table are visible and readable, particularly the upper section.</i> <p><i>The goal is to create a more legible, cleaner, and properly ordered Gantt chart that effectively communicates the project timeline, from initiation to completion, in a visually appealing manner".</i></p>
PS3	Prompt Designed for Monte Carlo Risk Analysis and Calculating a Probabilistic Duration for the Project	<p><i>"Given the complexity and uncertainty inherent in construction projects, including factors such as weather conditions, supply chain disruptions, and labor availability, we aim to estimate the probabilistic duration of a construction project using Monte Carlo simulation. To accomplish this, follow these steps:</i></p> <ol style="list-style-type: none"> <i>Assume Variability in Phase Durations: Model the duration (in days) of each project phase as a random variable using a triangular distribution, defined by optimistic, most likely, and pessimistic durations. Assume a 10% variability for the optimistic and pessimistic durations relative to the most likely duration for each phase.</i> <i>Monte Carlo Simulation Steps:</i> <ol style="list-style-type: none"> <i>Define Distributions for Each Phase: Assign triangular distributions for each project phase based on provided durations, incorporating the assumed variability.</i> <i>Simulate Total Project Duration: For each iteration of the simulation, randomly sample from each phase's distribution to calculate a total project duration (in days).</i> <i>Repeat and Aggregate Results: Conduct thousands of iterations (e.g., 10,000) to create a distribution of total project durations.</i> <i>Analyze Results: Determine key probabilistic metrics, such as the median duration and confidence intervals (e.g., the 95% confidence interval).</i> <i>Expected Outcomes: Provide a probabilistic estimate of the project's duration, including the median duration and a confidence interval that captures the range of potential outcomes given the inherent uncertainties.</i> <i>Visualization: Generate a distribution diagram that visually represents the outcomes of the Monte Carlo simulation, highlighting the median duration (in days) and the confidence interval to aid in understanding the variability and central tendency of the project durations.</i> <p><i>The objective is to perform a Monte Carlo simulation to offer a probabilistic estimate of the duration (in days) of a construction project, taking into account the uncertainties and variabilities typical of such projects. This analysis aims to provide a more nuanced understanding of project timelines, enabling better planning and risk management".</i></p>

It must be noted that these prompts align with the prompt optimization techniques introduced in the methodology as follows:

Prompt PS1:

- **Instructional Prompts:** The prompt clearly instructs ChatGPT on the format and content required (table details, specific columns, and calculations).
- **Constraint-based Prompts:** It specifies constraints such as excluding a particular phase and starting on a specific day.
- **Contextual Prompts:** Provides context about the project and what is excluded to generate coherent output.

Prompt PS2:

- **Instructional Prompts:** Provides detailed instructions on how to format and present the Gantt chart.
- **Constraint-based Prompts:** Specifies constraints like removing every other date and using a specific title and document format.
- **Reformulating and Rewriting Prompts:** Adjusts the original request to include detailed steps and formatting instructions.

Prompt PS3:

- **Instructional Prompts:** Provides a detailed set of instructions for the Monte Carlo simulation process.
- **Constraint-based Prompts:** Defines specific constraints such as using triangular distributions and a 10% variability.
- **Priming Prompts:** Seeds the model with initial context about the complexity and uncertainty in construction projects.
- **Breaking Down Complex Tasks:** Segments the task into manageable steps, improving precision and effectiveness.
- **Visualization Prompts:** Requests a visual representation of the simulation outcomes.

By applying these prompt optimization techniques, the prompts are structured to guide ChatGPT effectively, ensuring coherent, accurate, and contextually appropriate responses.

In the following, each case study will detail ChatGPT's generated output and the actual project timelines, analyzing the model's accuracy in schedule prediction and planning in the construction industry.

4.1. Case Study 1: Residential Building

The first case study is the construction of a 6000-square-foot custom home in Maryland, initiated on 5 June 2014, and culminating on 16 April 2015, with total duration of 315 days [27]. This project was planned using Microsoft Project scheduling software, detailing a series of critical construction phases from initial contracts and permit acquisitions to foundation laying, framing, and final inspections. It must be noted that any project timeline reflects the inherent variability in construction, influenced by factors such as site conditions, weather, material availability, and the intricacies of custom features.

4.1.1. Duration Estimation

The first task is to ask ChatGPT to create a schedule for the project by using the prompt provided in Table 1. The project constraints (phases excluded, potential for fast-tracking, weather considerations, etc.) are provided to ChatGPT, if provided by the reference. For the first case study, the construction process, from obtaining permits to completion, is defined by ChatGPT as: permitting, site preparation, foundation, framing, roofing, windows and doors, plumbing, electrical, HVAC, insulation, drywall, interior finishes, exterior finishes, and final inspections. The project schedule is outlined in a table format, considering the start date of 5 June 2014 (provided in the reference). It must be noted that the durations are estimated based on typical construction timelines, but actual durations can vary based on the specific details of the project, local regulations, weather conditions, and other unforeseen delays.

The total duration for this construction project, starting from 5 June 2014, and including all outlined phases, is estimated to be 294 days (Figure 7). This means the project is expected

to be completed by 26 March 2015. This schedule assumes continuous work without any days off and accounts for the linear and dependent nature of tasks. ChatGPT estimates that the project will take 294 days.

ID	Outline Number	Name	Duration (days)	Predecessors	Start Date	Finish Date
1	1	Obtaining Permits	30	None	06/05/2014	07/04/2014
2	2	Site Preparation	14	1	07/05/2014	07/18/2014
3	3	Foundation	21	2	07/19/2014	08/08/2014
4	4	Framing	35	3	08/09/2014	09/12/2014
5	5	Roofing	14	4	09/13/2014	09/26/2014
6	6	Windows and Doors	14	5	09/27/2014	10/10/2014
7	7	Plumbing	21	6	10/11/2014	10/31/2014
8	8	Electrical	21	6	10/11/2014	10/31/2014
9	9	HVAC	21	6	10/11/2014	10/31/2014
10	10	Insulation	7	7,8,9	11/01/2014	11/07/2014
11	11	Drywall	14	10	11/08/2014	11/21/2014
12	12	Interior Finishes	45	11	11/22/2014	01/05/2015
13	13	Exterior Finishes	30	12	01/06/2015	02/04/2015
14	14	Final Inspections	7	13	02/05/2015	02/11/2015

Figure 7. Schedule created by ChatGPT for Case 1.

4.1.2. Gantt Chart

The next task is to ask ChatGPT to create a Gantt chart for the project, detailing each phase of the project in a table. This Gantt chart is shown in Figure 8.

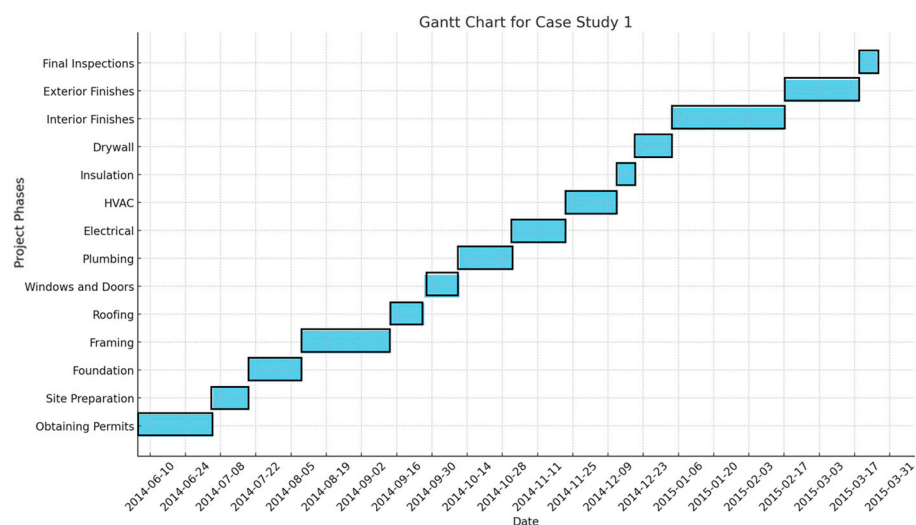


Figure 8. Gantt chart created by ChatGPT for Case 1.

4.1.3. Monte Carlo Simulation

The last step is Monte Carlo simulation, incorporating the variability and uncertainties typical of construction projects to provide a probabilistic estimate of the project duration. The results show that the median duration for the construction project is approximately 294 days, and the 95% confidence interval ranges from about 287 days to 301 days. This interval captures the range of potential outcomes, highlighting the impact of uncertainties such as weather conditions, supply chain disruptions, and labor availability on the project timeline.

The distribution diagram (Figure 9) visualizes the outcomes of the simulation, with the median duration in red and the confidence interval boundaries in green and blue.

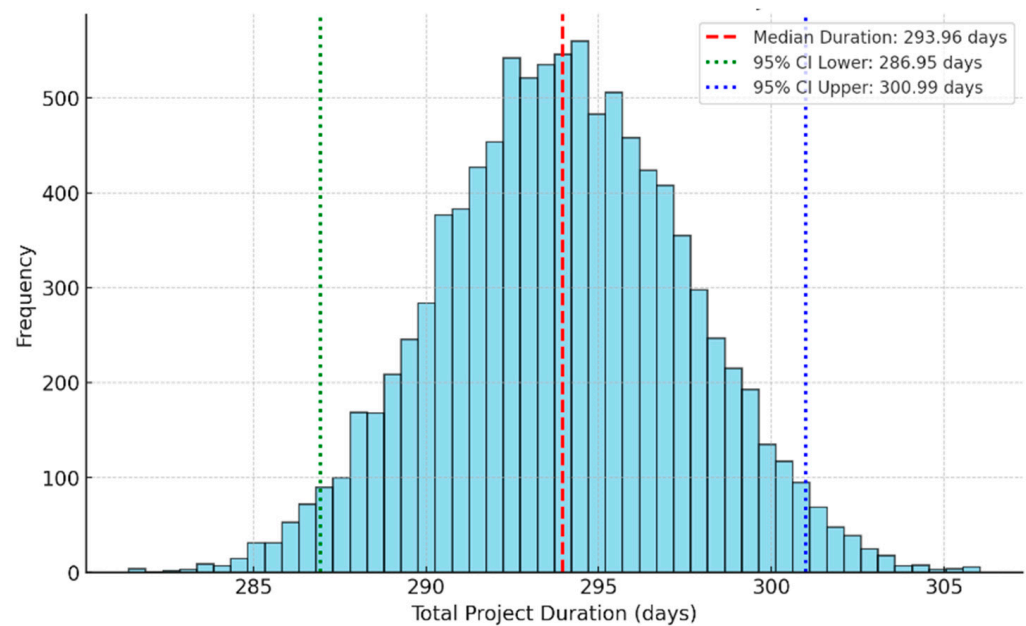


Figure 9. Monte Carlo project duration simulation created by ChatGPT for Case 1.

4.2. Case Study 2: Multipurpose Building

This study involves the construction of a shopping mall, starting on 15 June 2020, and ending on 12 November 2021 [28]. This project spans 496 days, starting with a pre-planning phase and moving through various critical stages including consultation, architectural design, detailed planning, execution, and finalization phases. Each task, such as feasibility testing, stakeholder analysis, and detailed estimation, is planned with specified start and finish dates, ensuring a comprehensive approach to managing the construction process efficiently. This case study showcases the complexity of large-scale construction projects and the importance of detailed scheduling to meet projected deadlines.

4.2.1. Duration Estimation

The first task, project tasks, and duration for each task is estimated by ChatGPT. Start and finish dates for each project are also estimated. ChatGPT is provided with the reference assumptions of minimal earthwork and expedited inspection and closing to provide a better estimation. ChatGPT estimates that the project will take 510 days (Figure 10).

ID	Outline Number	Name	Duration (Days)	Predecessors	Start Date	Finish Date
1	1	Pre-Planning and Design	60	None	15/06/2020	14/08/2020
2	2	Site Preparation	30	1	14/08/2020	13/09/2020
3	3	Foundation and Structural Framework	90	2	13/09/2020	12/12/2020
4	4	Exterior Construction	120	3	12/12/2020	11/04/2021
5	5	Interior Rough-in	90	4	11/04/2021	10/07/2021
6	6	Interior Finishing	120	5	10/07/2021	07/11/2021
7	7	Exterior Finishes and Landscaping	60	4	07/11/2021	06/01/2022
8	8	Specialty Installations	60	6,7	06/01/2022	07/03/2022
9	9	Inspections and Closing	30	8	07/03/2022	06/04/2022

Figure 10. Schedule created by ChatGPT for Case 2.

4.2.2. Gantt Chart

As the second task, a Gantt chart is created by ChatGPT to visually illustrate the schedule of the project (Figure 11).

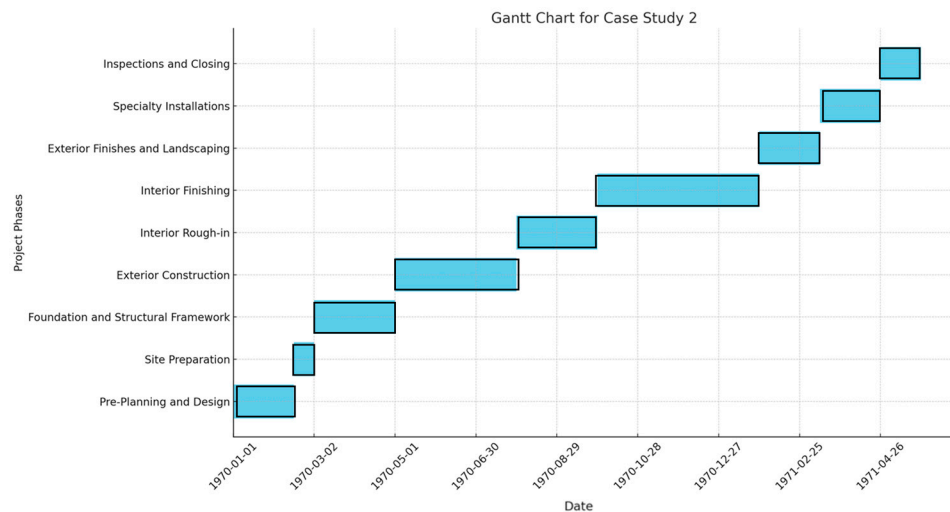


Figure 11. Gantt chart created by ChatGPT for Case 2.

4.2.3. Monte Carlo Simulation

The Monte Carlo simulation, considering the variability and uncertainties inherent in construction projects, is created as task three, to provide a probabilistic estimate of the project duration. The median duration, representing the central tendency of the project duration, is approximately 510 days (Figure 12). The 95% confidence interval ranges from approximately 495 days to 525 days, indicating that there is a 95% chance the project duration will fall within this range, given the assumed uncertainties.

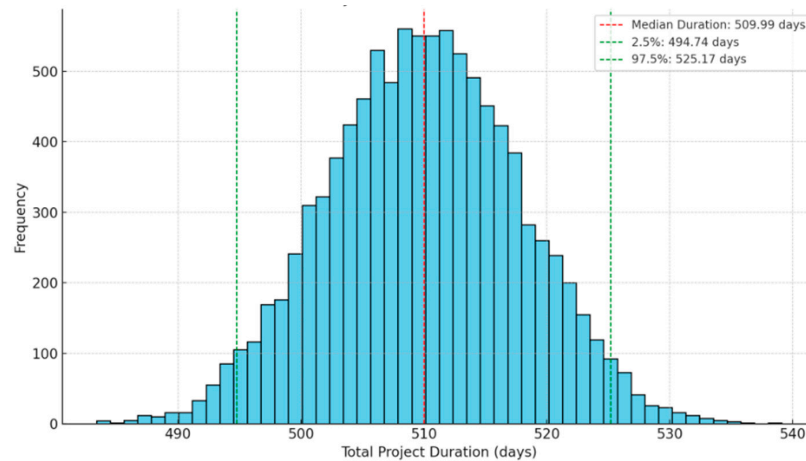


Figure 12. Monte Carlo project duration simulation created by ChatGPT for Case 2.

4.3. Case Study 3: Office Building Reconfiguration

The third case study entails the reconfiguration of a main office building and engineering and operations lobby, taking 404 days to complete [29]. This project, situated within the New Jersey Department of Transportation Headquarters, aims to revamp the main as well as engineering and operations lobbies to elevate safety and security. The enhancement includes the integration of turnstiles, the installation of slip-resistant flooring, and the construction of new security stations. Adjustments to windows and doors are anticipated to accommodate these updates. With a timeline of 404 days, this initiative not only seeks to modernize the infrastructure but also to significantly improve the security posture, ensuring a safer environment for both employees and visitors.

4.3.1. Duration Estimation

Initially, a schedule is created for the project by estimating the project phases, their duration, and precedence information (Figure 13). This case study was challenging, as there are lots of deliverables for the project. The model is created with the exact assumption that “[t]he building is divided into four wings per floor with about 310 occupants. The lobby reconfiguration will be confined to the first two floors of the main entrance” to provide the best result. ChatGPT estimates that the project will take 425 days.

ID	Outline Number	Name	Duration (Days)	Predecessors	Start Date	Finish Date
1	1	Initial Assessment	30	None	2024-04-01	2024-05-01
2	2	Design & Planning	90	1	2024-05-02	2024-07-31
3	3	Permitting	60	2	2024-08-01	2024-09-30
4	4	Contractor Selection	30	3	2024-10-01	2024-10-31
5	5	Lobby Reconfiguration Construction	180	4	2024-11-01	2025-04-30
6	6	Finishing Touches & Inspections	30	5	2025-05-01	2025-05-31

Figure 13. Schedule created by ChatGPT for Case 3.

4.3.2. Gantt Chart

Building on the estimated schedule, a Gantt chart is created for the project, illustrating the main phases and their start and finish dates (Figure 14). Since the main reference did not include any project start date, it is assumed that the project will start on April 1st, 2024.

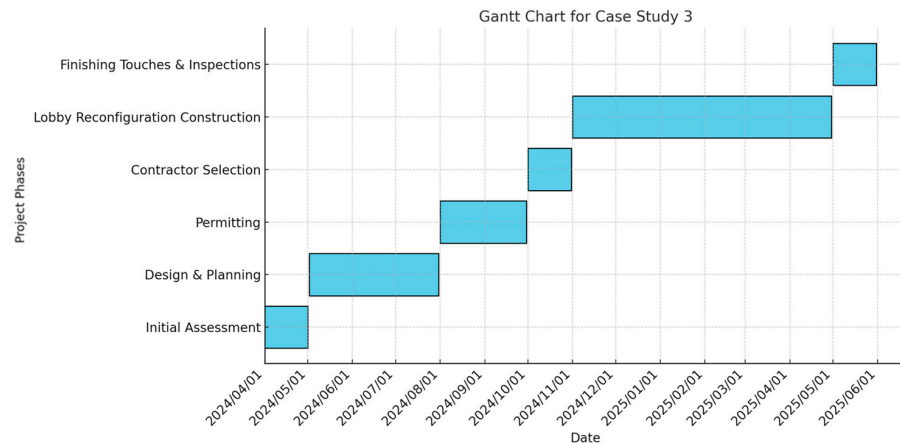


Figure 14. Gantt chart created by ChatGPT for Case 3.

4.3.3. Monte Carlo Simulation

The Monte Carlo simulation, conducted with 10,000 iterations to account for variability and uncertainty in construction project phase durations, provides a probabilistic estimate of the project’s duration. The median project duration is approximately 420 days (Figure 15). This value represents the central tendency of the project’s total duration, where half of the simulated outcomes are shorter, and half are longer. The simulation results suggest a 95% confidence interval for the total project duration ranging from approximately 403 days (2.5th percentile) to 437 days (97.5th percentile). This interval captures the range of potential outcomes, considering the inherent uncertainties and variabilities typical of construction projects.

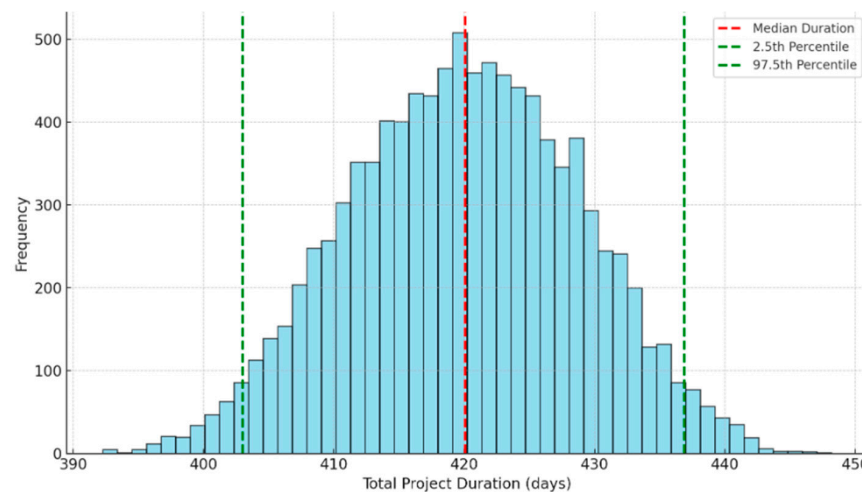


Figure 15. Monte Carlo project duration simulation created by ChatGPT for Case 3.

4.4. Case Study 4: Business Park Development

The next case study focuses on a business park in eastern Helsinki, consisting of a 14,500 m² office building with two independent sections and a parking hall, and constructed from 5 January 2004 to December 2005, taking 593 days [30].

4.4.1. Duration Estimation

Similar to other cases, the first task is to create a schedule for the project, by estimating different tasks and their duration. ChatGPT has assumed that the project “is composed of two sections, which can be built independently of each other and of parking hall below the main building”, and “[b]oth sections have six floors” to create better results. Also, the exclusion/inclusion criteria are defined as provided by the reference as “excluding

the ‘pre-design and design’ phases and only including the construction phase”. ChatGPT estimates that the project will take 555 days as it creates the schedule, shown in Figure 16.

ID	Outline Number	Name	Duration	Predecessors	Start Date	Finish Date
1	1	Site Preparation	30 days		05/01/2004	05/31/2004
2	2	Foundation and Basement Construction	60 days	1	05/31/2004	07/30/2004
3	3	Structural Framework	90 days	2	07/30/2004	10/28/2004
4	4	Roofing	30 days	3	10/28/2004	11/27/2004
5	5	Exterior Finishing	60 days	4	11/27/2004	01/26/2005
6	6	Interior Finishing	90 days	5	01/26/2005	04/26/2005
7	7	Mechanical and Electrical Installations	60 days	6	04/26/2005	06/25/2005
8	8	Parking Hall Construction	90 days	7	06/25/2005	09/23/2005
9	9	Landscaping and External Works	30 days	8	09/23/2005	10/23/2005
10	10	Final Inspection and Handover	15 days	9	10/23/2005	11/07/2005

Figure 16. Schedule created by ChatGPT for Case 4.

4.4.2. Gantt Chart

As the second task, a Gantt chart is created by ChatGPT to visually illustrate the schedule of the project (Figure 17).

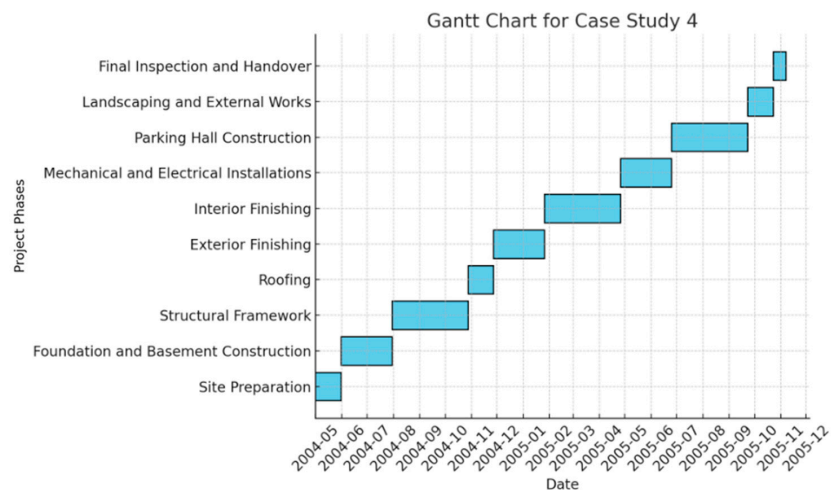


Figure 17. Gantt chart created by ChatGPT for Case 4.

4.4.3. Monte Carlo Simulation

The Monte Carlo simulation, considering the variability and uncertainty inherent in construction projects, provides a probabilistic estimate of the project’s duration. Based on 10,000 iterations, the median duration is approximately 555 days (Figure 18), and the 95% confidence range is between 539 days to 571 days. This analysis indicates that, with a high level of confidence, the project duration will likely fall within the calculated range, considering the uncertainties such as weather conditions, supply chain disruptions, and labor availability.

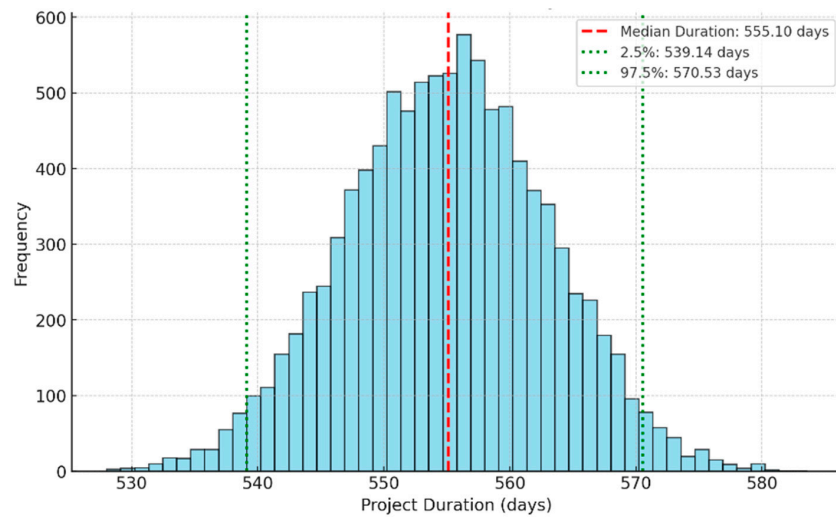


Figure 18. Monte Carlo project duration simulation created by ChatGPT for Case 4.

4.5. Case Study 5: Luxury Villa

The last case study is a residential project for a luxury villa with a construction timeline of 205 days [31]. There is no inclusion/exclusion criterion provided in the reference study.

4.5.1. Duration Estimation

A schedule is created by ChatGPT for this project, given the assumption that “landscaping and final inspections are excluded”. Main phases are determined by ChatGPT and a duration is assigned to each phase to ultimately estimate the project duration as 200 days (Figure 19).

ID	Outline Number	Name	Duration	Predecessors	Start Date	Finish Date
1	1	Site Preparation	14 days		2024-03-01	2024-03-15
2	2	Foundation	21 days	1	2024-03-15	2024-04-05
3	3	Framing	30 days	2	2024-04-05	2024-05-05
4	4	Roofing	15 days	3	2024-05-05	2024-05-20
5	5	Exterior Finishes	20 days	4	2024-05-20	2024-06-09
6	6	Interior Rough-in	25 days	5	2024-06-09	2024-07-04
7	7	Insulation and Drywall	20 days	6	2024-07-04	2024-07-24
8	8	Interior Finishes	30 days	7	2024-07-24	2024-08-23
9	9	Flooring	15 days	8	2024-08-23	2024-09-07
10	10	Utilities and Hookups	10 days	9	2024-09-07	2024-09-17

Figure 19. Schedule created by ChatGPT for Case 5.

4.5.2. Gantt Chart

A Gantt chart is created for the project given the estimated schedule (Figure 20).

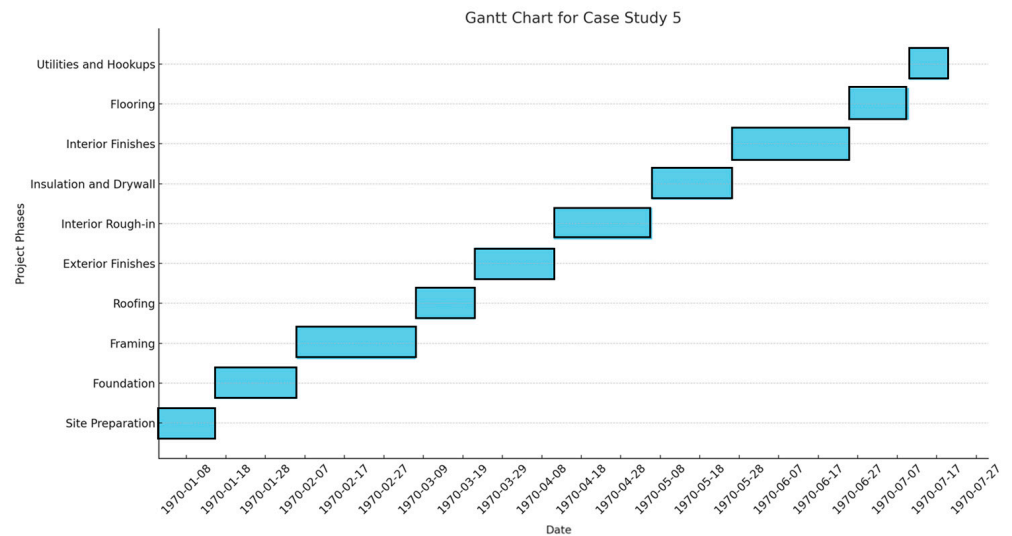


Figure 20. Gantt chart created by ChatGPT for Case 5.

4.5.3. Monte Carlo Simulation

The Monte Carlo simulation, considering the variability in the construction project's phase durations, has yielded a probabilistic estimate of the project duration. The median duration is approximately 200 days (Figure 21) and the range of potential outcomes, with a 95% probability that the project duration will fall within this range, is from 195 days to 205 days.

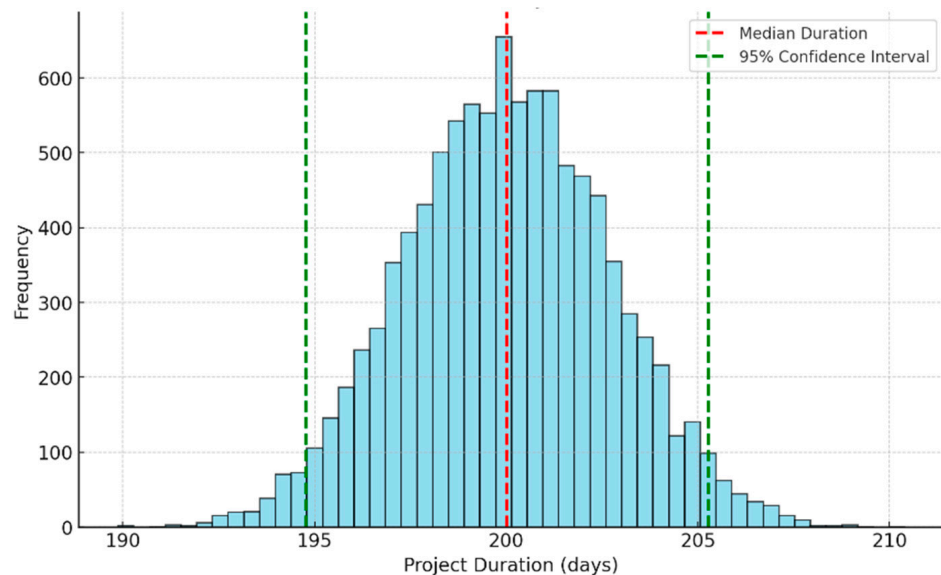


Figure 21. Monte Carlo project duration simulation created by ChatGPT for Case 5.

4.6. Discussion on Scheduling Case Studies

This section analyzes the effectiveness of ChatGPT in generating construction schedules for various project types. The comparison between the durations provided by reference, ChatGPT (Control), and ChatGPT (Prompt) along with their respective relative errors is summarized in Table 2. The table highlights that the durations generated using ChatGPT, where prompts are carefully engineered, are significantly closer to the reference durations

compared to those generated by ChatGPT without using any specific prompt engineering techniques (control). This indicates that the application of prompt engineering techniques has a considerable impact on improving the accuracy of ChatGPT's schedule predictions.

Table 2. Duration for each case study.

No.	Project Type	Duration Per Reference	Duration Per ChatGPT (Control)	Duration Per ChatGPT (Prompt)	Relative Error (Control)	Relative Error (Prompt)
1	Residential Building	315	450	294	0.43	0.07
2	Multipurpose Building	495	990	510	1.00	0.03
3	Office Building Reconfiguration	404	360	425	0.11	0.05
4	Business Park Development	593	1215	555	1.05	0.06
5	Luxury Villa	205	720	200	2.51	0.02

Overall, the results demonstrate that prompt engineering significantly enhances the precision of ChatGPT's output. For instance, the relative errors for ChatGPT (Prompt) across all project types are substantially lower than those for ChatGPT (Control), underscoring the importance of well-structured and contextually relevant prompts. This improvement is most notable in complex projects like multipurpose buildings and business park developments, where accurate scheduling is critical. These findings suggest that, by leveraging prompt engineering, practitioners in the AEC industry can harness the full potential of ChatGPT to generate more reliable and actionable project schedules, ultimately leading to better project management and execution.

5. Application 2: Construction Hazard Recognition

This section applies the designed prompt engineering methodology to five distinct construction hazard recognition case studies retrieved from OSHA. OSHA has provided a set of "Prevention Videos" simulating real incidents caused on construction job sites for training purposes [32]. A set of five common hazards are selected by the author, and screenshots are taken from the video, to be studied in this section. Each case involves utilizing ChatGPT4 to generate a list of potential hazards with OSHA references, which is then compared to the actual hazards documented for the case. The aim is to evaluate the capability of ChatGPT in detecting potential hazards on a job site.

- Case Study 1: Floor Opening.
- Case Study 2: Bridge Decking.
- Case Study 3: Swinging Crane.
- Case Study 4: Trenching.
- Case Study 5: Working Safely with Ladders Near Power Lines.

After several iterations using the methodology developed in this research, the following prompt (Table 3) is designed to generate the desired output based on an image provided depicting site conditions to ChatGPT4.

Similar to the Scheduling Prompts, the prompt engineering techniques explained in the methodology section are utilized to create the hazard recognition prompt as:

- Instructional Prompts: The prompt clearly provides step-by-step instructions for each task, guiding ChatGPT towards the desired output. It specifies the actions to be performed, such as identifying hazards, detailing OSHA procedures, creating an image, and generating a checklist.

- **Constraint-based Prompts:** The prompt sets explicit constraints by specifying the format (e.g., table format for the checklist, .doc document format for the final output) and the elements to be included in the image (guardrails, warning signs, personal protective equipment).
- **Contextual Prompts:** The prompt includes relevant information about the construction activity and the context of the tasks, helping ChatGPT generate coherent and contextually appropriate responses. For example, it mentions the specific activity and safety measures relevant to that activity.
- **Reformulating and Rewriting Prompts:** This prompt involves breaking down a complex task into smaller, more manageable tasks, guiding the model through each step to achieve accurate and comprehensive results.
- **Priming Prompts:** The prompt primes ChatGPT with specific instructions and expectations for each task, setting the stage for the model to build upon this information and generate detailed and relevant outputs.

Table 3. Hazard recognition prompts.

No.	Description	Prompt
PH1	Prompt Designed for Identifying the Hazards, OSHA Procedures, Safety Checklist, and Creating an Image with Corrective Actions	<p><i>“Given the attached image of a “Activity X” on a construction site, perform the following tasks:</i></p> <ol style="list-style-type: none"> <i>1. Identify Potential Hazards: Examine the image for any visible safety concerns related to the “Activity X” and list them.</i> <i>2. OSHA Safety Procedures: For each hazard identified, detail the corresponding safety procedures as per OSHA standards, including relevant references.</i> <i>3. Safety Implementation Image: Create a conceptual image that showcases a construction site with “Activity X”, incorporating all necessary safety measures identified in task 2. Include details such as guardrails, warning signs, and personal protective equipment.</i> <i>4. Safety Checklist: Generate a checklist in table format to review before starting work near a floor opening, covering all aspects of safety measures, training, and emergency preparedness.</i> <i>5. .doc Format: Lastly, provide the safety checklist from task 4 in a .doc document format”.</i>

By applying these prompt optimization techniques, these prompts are structured to ensure that ChatGPT can effectively identify hazards, outline OSHA procedures, create a conceptual image, and generate a safety checklist, all in a clear and precise manner.

An explanation on each case study, along with the outputs created by ChatGPT4, are demonstrated in the following subsection.

5.1. Case Study 1: Fall in Construction—Floor Opening Example

OSHA reports indicate that over 800 construction workers lose their lives annually on the job, with falls being the leading cause of these fatalities. Representing one-third of all construction-related deaths, these incidents occur quickly when workers are on elevated surfaces such as roofs, scaffolds, ladders, and bridges.

The first case (Figure 22) highlights the lethal risk of floor opening falls at construction sites and underscores the urgency for employers to ensure safer working conditions by adhering to mandated safety protocols and providing necessary protective gear. This example illustrates the tragic death of a worker who fell through an unprotected stairwell opening due to the lack of safety measures like guardrails or personal fall protection. By revisiting the sequence of events that led to this fatal accident, it is demonstrated that adherence to OSHA’s fall protection standards, such as installing guardrails or secure covers over openings can be lifesaving.



Figure 22. Floor opening simulation image by OSHA, Case Study 1.

5.1.1. Potential Hazards Identification

Upon examining the image provided, several safety concerns related to the floor opening are identified by ChatGPT4 as:

- **Unprotected Edges:** The floor opening does not seem to have any guardrails or covers, posing a fall hazard.
- **No Warning Signs:** There are no visible warning signs indicating the presence of a floor opening.
- **Lack of Personal Protective Equipment (PPE):** The individual in the image does not appear to be wearing any fall protection gear.
- **Tools Near Edge:** There are tools near the edge of the opening, which may present a tripping hazard or could fall through the opening.

5.1.2. OSHA Safety Procedures

For each hazard identified, the corresponding safety procedures as per OSHA standards are identified by ChatGPT4 as:

Unprotected Edges:

- OSHA requires that all floor holes into which a person can accidentally walk be guarded by either a standard railing with toeboard or a floor hole cover of standard strength and construction (Refer to OSHA 1910.28(b)(3)(i) [33]).

No Warning Signs:

- OSHA standard requires that the area should be marked with a warning sign or other indicators such as safety cones to warn of the open floor (Refer to OSHA 1926.502(f)(2) [33]).

Lack of Personal Protective Equipment (PPE):

- When work is performed at six feet or more above lower levels, the workers should use guardrails, safety nets, or personal fall arrest systems (Refer to OSHA 1926.501(b)(1) [33]).

Tools Near Edge:

- All materials and equipment should be kept far enough from the edge of the floor opening to prevent them from being knocked into the hole (Refer to OSHA 1926.501(b)(4) [33]).

5.1.3. Safety Implementation Image

ChatGPT4 created a conceptual image (Figure 23) that showcases a construction site with “Floor Openings”, incorporating all necessary safety measures as per the OSHA standards mentioned.

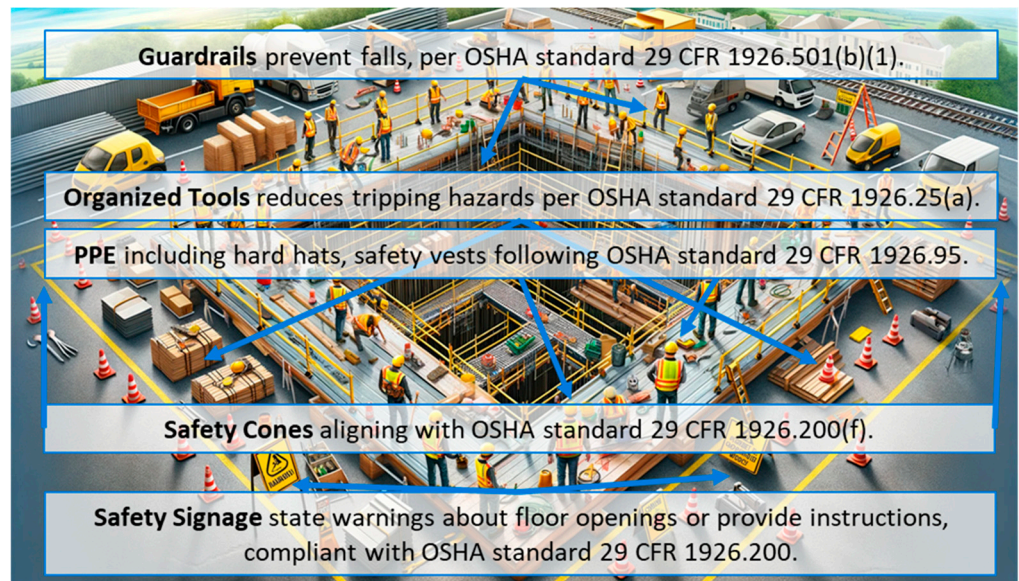


Figure 23. Safety implementation image created by ChatGPT for Case 1, Referring to [33].

5.1.4. Safety Checklist

Ultimately, a safety checklist is created by ChatGPT4, as shown in Figure 24.

No.	Item Description	Confirmed (Y/N)
1	Guardrails and toe-boards installed around floor openings	Y/N
2	Floor hole covers are secure and capable of supporting workers	Y/N
3	Area around the floor opening is clean, orderly, and free of tripping hazards	Y/N
4	Warning signs are posted around the floor opening	Y/N
5	Ladders are secured and equipped with self-closing safety gates	Y/N
6	Workers have received training on working around floor openings	Y/N
7	Workers are equipped with personal protective equipment (PPE)	Y/N
8	Emergency response plan is in place for fall incidents	Y/N

Figure 24. Safety checklist created by ChatGPT for Case 1.

5.2. Case Study 2: Fall in Construction—Bridge Decking Example

Since falls are the leading cause of fatal construction incidents, another example on bridge decking is selected (Figure 25). In this incident, two workers were engaged in installing deck pans on a bridge without personal fall protection or a safety net below, leading to a tragic incident where one worker fell 75 feet to her death after tripping and dislodging an unsecured deck pan. This example emphasizes the crucial role of fall protection, mandated by OSHA for work at heights exceeding 15 feet, in preventing such fatalities. Implementing a temporary horizontal lifeline or a safety net could have dramatically altered the outcome, illustrating how adherence to safety standards and the provision of proper fall protection equipment are vital measures that can prevent construction-related deaths.



Figure 25. Simulation image by OSHA, Case Study 2.

5.2.1. Potential Hazards

Upon examining the image provided, several safety concerns related to the bridge decking are identified by ChatGPT4 as:

- **Fall Hazard:** There is a significant drop visible next to the bridge decking, which presents a fall hazard.
- **Tripping Hazard:** Rebar protrusions can be tripping hazards.
- **Lack of Guardrails:** There are no guardrails on the open sides of the bridge decking, increasing the risk of falls.
- **Personal Protective Equipment (PPE):** The individual closest to the edge does not appear to be wearing a harness or any fall protection gear.

5.2.2. OSHA Safety Procedures:

For each of the identified hazards, OSHA has set standards to mitigate the risks. ChatGPT identifies these procedures as:

Fall Protection:

- Where there is a risk of falling six feet or more, OSHA requires employers to provide fall protection systems, such as guardrails, safety nets, or personal fall arrest systems (Refer to 29 CFR 1926.501 [33]).

Guardrail Systems:

- When work is being performed on a surface with unprotected sides and edges six feet above lower levels, guardrail systems should be installed (Refer to 29 CFR 1926.502(b) [33]).

Control of Hazardous Energy (Lockout/Tagout):

- Proper procedures should be implemented to ensure that equipment is de-energized and cannot be started up unexpectedly, which is relevant if any machinery is being used on the bridge decking (Refer to 29 CFR 1910.147 [33]).

Safety Training and Education:

- Employers must instruct each employee in the recognition and avoidance of unsafe conditions and the regulations applicable to their work environment to control or eliminate any hazards or other exposure to illness or injury (Refer to 29 CFR 1926.21(b)(2) [33]).

5.2.3. Safety Implementation Image

ChatGPT4 created a conceptual image (Figure 26) that showcases a construction site with “Bridge Decking”, incorporating all necessary safety measures as per the OSHA standards mentioned.



Figure 26. Safety implementation image created by ChatGPT for Case 2, Referring to [33].

5.2.4. Safety Checklist

Ultimately, a safety checklist is created by ChatGPT4, as shown in Figure 27.

No.	Item Description	Confirmed (Y/N)
1	Guardrail and edge protection systems are in place	
2	Personal fall arrest systems and training provided	
3	Decking gaps are covered or protected	
4	Secure access points to the bridge	
5	Install proper hazard signage	
6	Conduct safety training for working at heights	
7	Availability and use of PPE	
8	Emergency and rescue procedures established	
9	Regular equipment and safety systems inspections	
10	Work area free from tripping hazards and debris	
11	Proper storage and securing of materials/tools	
12	Electrical equipment checked for water safety	
13	Installation of safety nets, if applicable	
14	Adequate lighting for all work areas	
15	Monitoring of weather conditions	

Figure 27. Safety checklist created by ChatGPT for Case 2.

5.3. Case Study 3: Struck-By Accidents in Construction—Swinging Crane Example

The third simulated case highlights the potential struck-by hazards in construction sites. In this example (Figure 28), a truck driver was fatally crushed by a crane’s counterweight after inadvertently entering the crane’s swing radius. This case illustrates the lethal risk of being struck by heavy equipment, a leading cause of injury and the second most

common cause of death among construction workers, with over 150 fatalities per year. The incident reveals a critical oversight in workplace safety protocols, emphasizing the need for strict enforcement of secured zones around heavy machinery. Implementing simple yet effective measures, such as barricading the crane's swing radius and ensuring clear communication among workers, could prevent such devastating accidents.



Figure 28. Simulation image by OSHA, Case Study 3.

5.3.1. Identified Hazards

Upon examining the image provided, several safety concerns related to the swinging crane are identified by ChatGPT4 as:

- **Proximity to Power Lines:** The crane appears to be operating near a building that may have power lines nearby, which is a common hazard.
- **Overloading:** Without specific information, it is impossible to tell if the crane is properly loaded, but overloading is a risk that can lead to tipping.
- **Unsecured Load:** The load being lifted must be properly secured to prevent it from slipping.
- **Operator Visibility:** It is unclear if the crane operator has a clear line of sight for the entire path of the crane and load.
- **Worker Proximity:** There is a worker on the ground close to the operating crane which poses a risk of being struck by the load or the crane.
- **Stability of Crane:** The crane needs to be on stable ground and properly balanced to avoid tipping.

5.3.2. OSHA Safety Procedures:

For each of the identified hazards, OSHA has set standards to mitigate the risks. ChatGPT identifies these procedures as:

Proximity to Power Lines:

- Maintain a safe working clearance from all power lines (at least 10 feet) (Refer to 29 CFR 1926.1408 [33]).
- De-energize and ground power lines when working in close proximity where possible (Refer to 29 CFR 1926.1408 [33]).
- Use a dedicated spotter to ensure the crane stays a safe distance away from power lines (Refer to 29 CFR 1926.1408 [33]).

Overloading:

- Follow the crane's load chart for the maximum load capacity (Refer to 29 CFR 1926.1438 [33]).

- Perform pre-operation inspection and planning to ensure loads do not exceed the crane's rated capacity (Refer to 29 CFR 1926.1438 [33]).
- Unsecured Load:
 - Use proper rigging techniques to secure all loads (Refer to 29 CFR 1926.251 [33]).
 - Inspect rigging equipment before use for any wear or damage (Refer to 29 CFR 1926.251 [33]).

Operator Visibility:

- Ensure that the operator has a clear view of the crane's path and load movement. If not, use signal persons (Refer to 29 CFR 1926.1428 [33]).
- Use cameras or other aids if the operator's view is obstructed (Refer to 29 CFR 1926.1428 [33]).

Worker Proximity:

- Establish a perimeter to keep workers at a safe distance from the crane's operation (Refer to 29 CFR 1926.1406 [33]).
- Use barriers or warning signs to prevent workers from entering the swing radius (Refer to 29 CFR 1926.1406 [33]).

Stability of Crane:

- Conduct ground surveys to ensure the stability of the surface on which the crane will operate (Refer to 29 CFR 1926.1402 [33]).
- Use outriggers and stabilizers as per manufacturer's recommendations (Refer to 29 CFR 1926.1402 [33]).

5.3.3. Safety Implementation Image

ChatGPT4 created a conceptual image (Figure 29) that showcases a construction site with "Swinging Cranes", incorporating all necessary safety measures as per the OSHA standards mentioned.

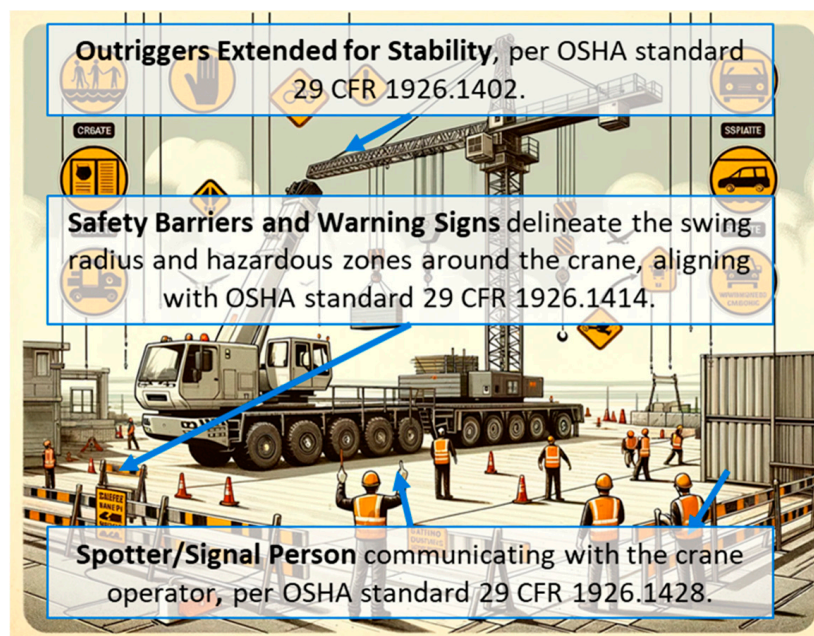


Figure 29. Safety implementation image created by ChatGPT for Case 3, Referring to [33].

5.3.4. Safety Checklist

Ultimately, a safety checklist is created by ChatGPT4, as shown in Figure 30.

Safety Checklist Item	Description	Completed (Y/N)
Load Rigging Inspection	Confirm that all loads are properly rigged as per OSHA 1926.251.	
Clearance Verification	Ensure there is sufficient clearance around the crane's operation as per OSHA 1926.1412(d).	
Barricade Establishment	Set up barricades around the crane's swing radius according to OSHA 1926.1424(a).	
Worker Positioning	Verify that no workers are within the crane's operating area in accordance with OSHA 1926.1425(c).	
Overhead Load Safety	Confirm the crane is not operating above workers unless specific conditions are met (OSHA 1926.1425(b)).	
Load Weight Check	Check that the crane is not overloaded and adheres to OSHA 1926.1417(a).	
PPE Distribution	Ensure all workers have and use proper personal protective equipment.	
Emergency Procedures Review	Review emergency preparedness plans and make sure all workers are trained on them.	
Floor Opening Guarding	Confirm that all floor openings are guarded by railings or covers as per OSHA standards.	
Signage and Signals	Make sure appropriate signs and hand signals are in place for crane and ground operations.	

Figure 30. Safety checklist created by ChatGPT for Case 3 [33].

5.4. Case Study 4: Excavations in Construction—Trenching Example

The fourth example is on trenching. Trenching work, known for its high risk, claims approximately 40 lives annually due to cave-ins. In this example (Figure 31) involving five workers in a trench without protective measures, three workers tragically suffocated due to the collapse. This incident emphasizes the critical need for compliance with OSHA standards, including proper trench protections like sloping, shoring, or shielding, to prevent such fatalities.



Figure 31. Simulation image by OSHA, Case Study 4.

OSHA mandates stringent rules for trench safety, emphasizing the weight of soil and the necessity of protective systems for trenches deeper than 5 feet or showing cave-in risks. Soil is classified from stable rock to type C, with each type requiring specific protective measures. Competent in-person inspections, soil testing, and the use of sloping, shoring, or shielding systems are essential to ensure worker safety in trenching operations, demonstrating the complexity and importance of adhering to safety protocols to prevent trench-related incidents.

5.4.1. Identified Hazards

Upon examining the image provided, several safety concerns related to trenching are identified by ChatGPT4 as:

- **Protective Systems:** Lack of protective systems to prevent cave-ins.
- **Safe Egress:** No visible means of safe egress such as ladders or steps within 25 feet of workers.
- **Hazardous Atmospheres:** The trench appears to be deeper than 4 feet, and there are no visible protective measures against potential hazardous atmospheres.
- **Barricades and Warning Signs:** There are no barricades or warning signs indicating the presence of a trench, which can pose a risk to the workers and the public.

5.4.2. OSHA Safety Procedures

For each of the identified hazards, OSHA has set standards to mitigate the risks. ChatGPT identifies these procedures as:

Protective Systems:

- For trenches deeper than 5 feet, OSHA requires a protective system unless the excavation is made entirely in stable rock. Options include sloping, shoring, benching, or using a trench box (Refer to 29 CFR 1926.652(a)(1) [33]).

Safe Egress:

- OSHA mandates that there must be safe means of egress, such as ladders or steps, within 25 lateral feet of workers in trenches 4 feet or deeper (Refer to 29 CFR 1926.651(c)(2) [33]).

Hazardous Atmospheres:

- When there is a potential for a hazardous atmosphere in trenches deeper than 4 feet, OSHA requires testing before entry and as needed throughout the work, with emergency rescue equipment available where hazardous conditions exist (Refer to 29 CFR 1926.651(g) [33]).

Barricades and Warning Signs:

- OSHA standards necessitate that barricades, hand or mechanical signals, or stop logs be used if there is a risk of equipment falling into or operating near the trench. Warning vests or other highly visible clothing are required for workers exposed to public vehicular traffic (Refer to 29 CFR 1926.651(k) [33]).

5.4.3. Safety Implementation Image

ChatGPT4 created a conceptual image (Figure 32) that showcases a construction site with "Trenching", incorporating all necessary safety measures as per the OSHA standards mentioned.

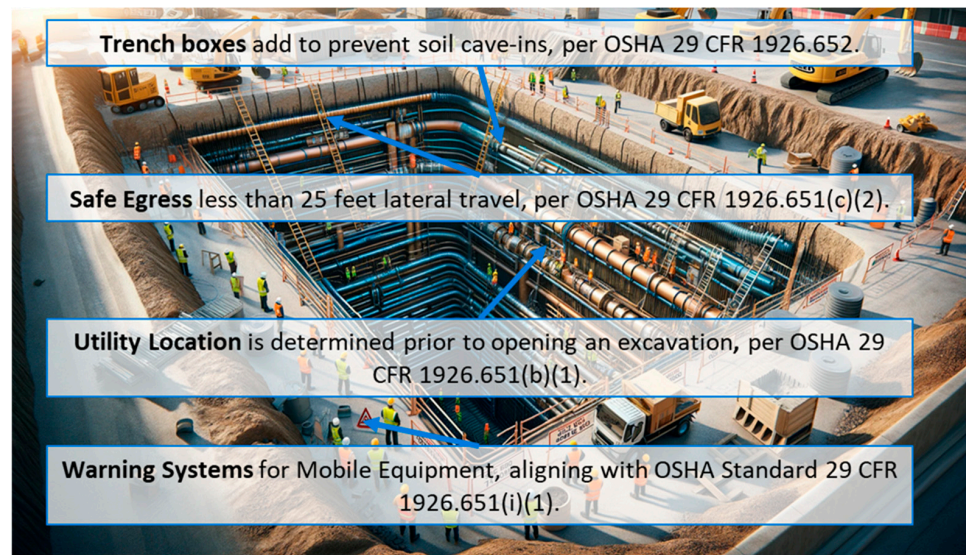


Figure 32. Safety implementation image created by ChatGPT for Case 4, Referring to [33].

5.4.4. Safety Checklist

Ultimately, a safety checklist is created by ChatGPT4, as shown in Figure 33.

No.	Item Description	Confirmed (Y/N)
1	Protective system in place for trench (e.g., trench box)	
2	Ladder or other safe means of egress within 25 feet of workers	
3	Hazardous atmosphere testing performed if necessary	
4	Emergency rescue equipment available on site	
5	Barricades and warning signs in place around trench perimeter	
6	High-visibility vests and hard hats worn by workers	
7	Daily inspection of trench conditions by a competent person	
8	Workers trained on trench safety and emergency procedures	
9	No equipment closer to the trench than two feet, unless the trench is properly supported	
10	Utilities (electric, gas, water, etc.) located and marked	

Figure 33. Safety checklist created by ChatGPT for Case 4.

5.5. Case Study 5: Electrocution-Working Safely with Ladders near Power Lines Example

The last case studies an electrocution incident, involving an employee working with a ladder near a power line. In this tragic incident of workplace electrocution (Figure 34), a construction worker was fatally electrocuted after his metal ladder made contact with overhead power lines during a window caulking task at a three-story townhouse. The incident, a direct result of inadequate safety measures and lack of proper equipment, underscores the critical need for comprehensive training, careful pre-job planning, and the adoption of appropriate precautions such as using non-conductive ladders and maintaining safe distances from power lines.



Figure 34. Simulation image by OSHA, Case Study 5.

5.5.1. Identified Hazards

Upon examining the image provided, several safety concerns related to electrocution are identified by ChatGPT4 as:

- **Proximity to Power Lines:** The ladder is placed very close to the power lines, which poses a risk of electrocution.
- **Ladder Angle:** It appears that the ladder may not be at the correct angle, which should be approximately 75 degrees from the ground for optimal stability.
- **Ladder Base Stability:** The base of the ladder needs to be on solid, level ground to prevent slipping or tipping.
- **Personal Protective Equipment:** It is not clear if the worker is wearing all the necessary personal protective equipment (PPE) for this task.

5.5.2. OSHA Safety Procedures

For each of the identified hazards, OSHA has set standards to mitigate the risks. ChatGPT identifies these procedures as:

Proximity to Power Lines:

- Maintain a minimum distance of 10 feet from overhead power lines (Refer to 29 CFR 1926.1408 [33]).
- Use a non-conductive ladder (e.g., fiberglass) when working near power lines (Refer to 29 CFR 1926.1408 [33]).
- De-energize and ground power lines when work is being performed near them, if possible (Refer to 29 CFR 1926.1408 [33]).

Ladder Angle:

- Use the 4-to-1 rule: for every 4 feet of ladder height, the base should be 1 foot away from the wall (Refer to 29 CFR 1926.1053(b)(1) [33]).
- Ladders should be used at an angle where the horizontal distance from the top support to the foot of the ladder is about 1/4 of the working length of the ladder (Refer to 29 CFR 1926.1053(b)(5) [33]).

Ladder Base Stability:

- Place the ladder on stable and level ground (Refer to 29 CFR 1926.1053(b)(5)(i) [33]).
- Secure the ladder at the base and top to prevent displacement (Refer to 29 CFR 1926.1053(b)(5)(ii) [33]).
- Extend the ladder 3 feet over the landing surface if you need to step onto another surface (Refer to 29 CFR 1926.1053(b)(16) [33]).

Personal Protective Equipment:

- Workers should wear rubber-soled, non-slip shoes (Refer to 29 CFR 1926.1053(b)(15) [33]).
- When working near electricity, workers should wear protective clothing and equipment to minimize the risk of electric shock (Refer to 29 CFR 1926.97 [33]).

5.5.3. Safety Implementation Image

ChatGPT4 created a conceptual image (Figure 35) that showcases a construction site with “Working Safely with Ladders Near Power Lines”, incorporating all necessary safety measures as per the OSHA standards mentioned.

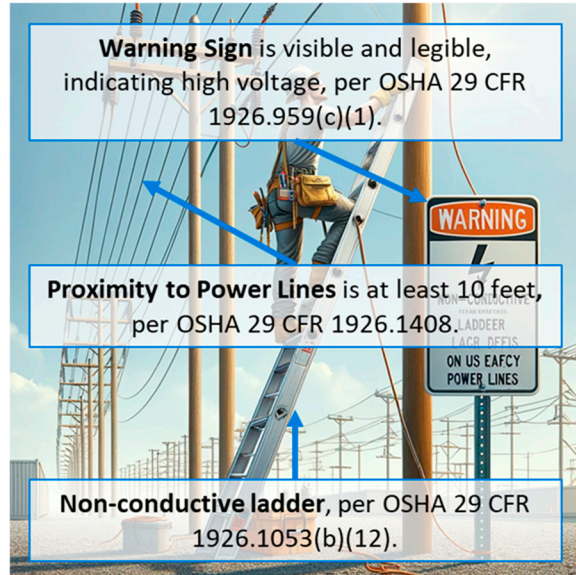


Figure 35. Safety implementation image created by ChatGPT for Case 5, Referring to [33].

5.5.4. Safety Checklist

Ultimately, a safety checklist is created by ChatGPT4, as shown in Figure 36.

No.	Item Description	Confirmed (Y/N)
1	Minimum 10-foot distance from power lines maintained	
2	Non-conductive ladder used	
3	Ladder angle follows the 4-to-1 rule	
4	Ladder base is stable and level	
5	Ladder extends 3 feet above the landing surface	
6	Ladder is secured at both the base and top	
7	Worker is wearing rubber-soled, non-slip shoes	
8	Worker is wearing all required PPE for electrical safety	
9	Warning signs for power lines are in place	
10	Emergency procedures for electrical contact are known to workers	
11	First aid kit and fire extinguisher are easily accessible	
12	Workers have received training on working safely with ladders	
13	Daily safety inspection of ladder and equipment completed	
14	Guardrails are installed on the roof where necessary	
15	All workers are aware of the location of the nearest hospital	

Figure 36. Safety checklist created by ChatGPT for Case 5.

5.6. Discussion on Hazard Recognition Case Studies

This section analyzes the effectiveness of ChatGPT in identifying hazards for various project types. The comparison between the hazards identified by reference, ChatGPT (Control), and ChatGPT (Prompt), along with their respective relative errors, is summarized in Table 4. The table highlights that the hazards identified using ChatGPT with carefully engineered prompts are significantly closer to the reference hazards compared to those identified by ChatGPT without using any specific prompt engineering techniques (control). This indicates that the application of prompt engineering techniques has a considerable impact on improving the accuracy of ChatGPT's hazard identification.

Table 4. Results for hazard recognition.

No.	Project Type	Hazards Per Reference	Hazards Per ChatGPT (Control)	Hazards Per ChatGPT (Prompt)	Relative Error (Control)	Relative Error (Prompt)
1	Floor Opening	4	3	4	0.25	0.00
2	Bridge Decking	4	3	4	0.25	0.00
3	Swinging Crane	6	5	6	0.17	0.00
4	Trenching	5	4	4	0.20	0.20
5	Working Safely with Ladders	4	4	4	0.00	0.00

Overall, the results demonstrate that prompt engineering significantly enhances the precision of ChatGPT's output. For instance, the relative errors for ChatGPT (Prompt) across most project types are substantially lower than those for ChatGPT (Control), underscoring the importance of well-structured and contextually relevant prompts. This improvement is most notable in tasks like identifying hazards in swinging crane operations, where accurate hazard identification is critical. However, it is worth noting that in more complex tasks like trenching, the relative error for ChatGPT (Prompt) was higher, indicating that there are still challenges to address in optimizing prompts for certain complex scenarios. These findings suggest that, by leveraging prompt engineering, practitioners in the AEC industry can harness the full potential of ChatGPT to generate more reliable and actionable hazard identifications, ultimately leading to better safety management and execution.

6. Conclusions

Generative AI and ChatGPT hold considerable promise for aiding the AEC industry across various applications, such as construction scheduling and hazard recognition. Additionally, there is potential for enhancing information retrieval from BIM and similar platforms. The importance of prompt engineering cannot be overstated in ensuring ChatGPT offers valuable and accurate insights. By utilizing the techniques and methodologies detailed in this paper, users can achieve results that are not only more accurate and relevant but also demonstrate greater creativity from this powerful AI language model.

The results of this study highlight the significant impact of prompt engineering on the accuracy of ChatGPT's outputs in both construction scheduling and hazard recognition. As shown in Table 2 for construction scheduling, the relative errors for ChatGPT (Prompt) are substantially lower than those for ChatGPT (Control). For example, in the case of the multipurpose building, the relative error for ChatGPT (Prompt) was 0.03 compared to 1.00 for ChatGPT (Control). Similarly, for the business park development, the relative error for ChatGPT (Prompt) was 0.06 compared to 1.05 for ChatGPT (Control). This demonstrates that prompt engineering techniques significantly improve the accuracy of ChatGPT's schedule predictions, making them much closer to the reference durations.

For hazard recognition, as detailed in Table 4, the hazards identified using ChatGPT with carefully engineered prompts are significantly closer to the reference hazards com-

pared to those identified by ChatGPT without using any specific prompt engineering techniques (Control). For instance, the relative errors for ChatGPT (Prompt) across most project types are substantially lower than those for ChatGPT (Control). For floor openings, bridge decking, and swinging cranes, the relative error for ChatGPT (Prompt) was 0.00, compared to 0.25 and 0.17 for ChatGPT (Control). However, in the more complex task of trenching, the relative error for ChatGPT (Prompt) was higher at 0.20, indicating that there are still challenges to address in optimizing prompts for certain scenarios.

While this study primarily addresses the specific cases of construction scheduling and hazard recognition, the principles of prompt engineering possess broader applicability across various domains within the AEC industry. The optimized prompts developed for these tasks can function as adaptable templates, suitable for a wide range of applications such as project cost estimation, resource allocation, and quality control. For instance, in project cost estimation, prompts could be customized to incorporate constraints related to budget categories or cost limits, while in quality control they might be tailored to include specific standards or regulatory requirements.

Overall, the study demonstrates that prompt engineering significantly enhances the precision of ChatGPT's output, making it a more reliable tool for both construction scheduling and hazard identification in the AEC industry. This improvement is particularly notable in tasks like identifying hazards in swinging crane operations and generating schedules for complex projects like multipurpose buildings and business park developments, where accurate identification and scheduling are critical. These findings suggest that by leveraging prompt engineering, practitioners in the AEC industry can harness the full potential of ChatGPT to generate more reliable and actionable insights, ultimately leading to better project and safety management.

6.1. Challenges and Limitations

While ChatGPT holds considerable promise for transforming the AEC industry, several practical challenges must be tackled to ensure its successful implementation in real-world scenarios. One such challenge involves integrating ChatGPT seamlessly with existing BIM across various stages, from planning to execution. Ensuring AI compatibility with these systems is essential for its broader application.

Another key challenge relates to data privacy. As AI becomes more embedded in the AEC industry, addressing the regulatory and ethical implications surrounding its use is crucial. One critical issue is liability—if an AI tool like ChatGPT makes an error, it is important to determine who is accountable and establish protocols for managing such situations. This highlights the need for clear guidelines and legal frameworks to mitigate AI-related risks in construction projects.

Additionally, ethical considerations around data usage are paramount. Since construction projects often involve sensitive information, it is crucial that AI systems safeguard this data responsibly, maintaining privacy and adhering to legal standards. Furthermore, as the industry evolves, AI tools must be continually updated to align with the latest standards and regulations. A thorough examination of these regulatory and ethical aspects will better equip the AEC industry to integrate AI responsibly and effectively.

It is also essential to recognize that AI should be used to enhance human understanding rather than foster dependency or oversimplification. While tools like ChatGPT can significantly improve efficiency in the AEC industry, human oversight and collaboration remain indispensable. AI should complement, not replace, human expertise. Given the complexity of the construction industry, which often demands complex decision-making, integrating AI-driven insights with human experience is critical. Future research should explore how AI can work alongside AEC professionals, supporting and amplifying their judgment and skills. This collaborative approach ensures that AI tools are used thoughtfully, adding value to the decision-making process while preserving the crucial human element in construction projects.

While the practical challenges outlined above must be addressed for successful AI integration, there are also inherent limitations within the current AI models that need to be acknowledged. For instance, there is a risk of inaccuracies in ChatGPT responses that may not always be readily apparent, highlighting the need for ensuring the reliability of this tool. One of the critical challenges in applying AI models like ChatGPT within the AEC industry is the potential for biases inherent in the model's training data, which can result in skewed or incomplete hazard identification. Given that the training data for LLMs may not comprehensively encompass the full spectrum of industry-specific hazard standards, there is a significant risk that the outputs generated by these models may not fully align with the latest safety regulations or may include inaccuracies, commonly referred to as 'hallucinations'.

Last, but not least, the dynamic nature of construction safety regulations poses an additional challenge, as it is improbable that any single LLM will remain fully aligned with these evolving standards without regular updates and fine-tuning. Consequently, while this study highlights the potential of ChatGPT in enhancing hazard recognition, it is imperative to caution that AI-generated outputs should be consistently cross-referenced with the most recent industry guidelines and expert insights to ensure their accuracy and relevance.

6.2. Future Research Directions

Future studies will expand the evaluation of the ChatGPT-based tool beyond the initial accuracy rate to include detailed assessments across four key dimensions: validity, explainability, usefulness, and comprehensiveness, using a Likert scale from 1 (low satisfaction) to 4 (high satisfaction). The validity of the experimental setup will also be evaluated to ensure it accurately measures the intended variables and aligns with research objectives. Explainability will be analyzed to assess the clarity and comprehensibility of ChatGPT's outputs, enhancing user understanding and trust in the tool's logic. The usefulness of ChatGPT's responses in real-world construction management scenarios will be investigated, determining their relevance and value in practical decision-making processes. Lastly, the comprehensiveness of the responses will be evaluated to ensure they cover all necessary aspects of the prompts, providing thorough and complete task handling. Through these extensive evaluations, the utility of ChatGPT will be refined and enhanced, making it a more effective and reliable tool in construction management and other related fields.

As this research is among the first studies on the application of ChatGPT in construction and is in its early stages, further refinement and continual testing and validation are required. As AI and NLP evolve, new avenues of research and application will unfold in prompt engineering tailored for the AEC sector. Potential directions for investigation include refining prompt strategies, incorporating external resources and APIs, and crafting interactive, multi-turn conversational systems. These advancements will lay the groundwork for AI language models such as ChatGPT to extend their versatility and enhance their utility across a wide range of applications.

Funding: This research received no external funding.

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the author on request.

Conflicts of Interest: The author declares no conflicts of interest.

References

1. Lubiana, T.; Lopes, R.; Medeiros, P.; Silva, J.C.; Goncalves, A.N.A.; Maracaja-Coutinho, V.; Nakaya, H.I. Ten quick tips for harnessing the power of ChatGPT in computational biology. *PLoS Comput. Biol.* **2023**, *19*, e1011319. [[CrossRef](#)] [[PubMed](#)]
2. Chung, S.; Moon, S.; Kim, J.; Kim, J.; Lim, S.; Chi, S. Comparing natural language processing (NLP) applications in construction and computer science using preferred reporting items for systematic reviews (PRISMA). *Autom. Constr.* **2023**, *154*, 105020. [[CrossRef](#)]
3. Dale, R. GPT-3: What's it good for? *Nat. Lang. Eng.* **2021**, *27*, 113–118. [[CrossRef](#)]

4. Budzianowski, P.; Vulić, I. Hello, it's GPT-2—How can I help you? towards the use of pretrained language models for task-oriented dialogue systems. *arXiv* **2019**, arXiv:1907.05774.
5. Saka, A.; Taiwo, R.; Saka, N.; Salami, B.A.; Ajayi, S.; Akande, K.; Kazemi, H. GPT models in construction industry: Opportunities, limitations, and a use case validation. *Dev. Built Environ.* **2023**, *17*, 100300. [[CrossRef](#)]
6. Radford, A.; Narasimhan, K.; Salimans, T.; Sutskever, I. Improving Language Understanding by Generative Pre-Training. 2018. Available online: https://cdn.openai.com/research-covers/language-unsupervised/language_understanding_paper.pdf (accessed on 1 June 2024).
7. Zheng, J.; Fischer, M. BIM-GPT: A prompt-based virtual Assistant framework for BIM information retrieval. *arXiv* **2023**, arXiv:2304.09333.
8. You, H.; Ye, Y.; Zhou, T.; Zhu, Q.; Du, J. Robot-Enabled Construction Assembly with Automated Sequence Planning Based on ChatGPT: RoboGPT. *Buildings* **2023**, *13*, 1772. [[CrossRef](#)]
9. Uddin, S.M.J.; Albert, A.; Ovid, A.; Alsharef, A. Leveraging ChatGPT to Aid Construction Hazard Recognition and Support Safety Education and Training. *Sustainability* **2023**, *15*, 7121. [[CrossRef](#)]
10. Prieto, S.A.; Mengiste, E.T.; de Soto, B.G. Investigating the Use of ChatGPT for the Scheduling of Construction Projects. *Buildings* **2023**, *13*, 857. [[CrossRef](#)]
11. Amer, F.; Jung, Y.; Golparvar-Fard, M. Transformer machine learning language model for auto-alignment of long-term and short-term plans in construction. *Autom. Constr.* **2021**, *132*, 103929. [[CrossRef](#)]
12. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, L.; Polosukhin, I. Attention Is All You Need. In Proceedings of the 31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA, 4–9 December 2017.
13. Kotei, E.; Thirunavukarasu, R. A Systematic Review of Transformer-Based Pre-Trained Language Models through Self-Supervised Learning. *Information* **2023**, *14*, 187. [[CrossRef](#)]
14. Levy, S.M. *Construction Process Planning and Management—Chapter 5—Preparing the Bid Documents*; Elsevier: Amsterdam, The Netherlands, 2010; pp. 113–135.
15. Guo, K.; Zhang, L. Multi-objective optimization for improved project management: Current status and future directions. *Autom. Constr.* **2022**, *139*, 104256. [[CrossRef](#)]
16. Lu, Y.; Sood, T.; Chang, R.; Liao, L. Factors impacting integrated design process of net zero energy buildings: An integrated framework. *Int. J. Constr. Manag.* **2020**, *22*, 1700–1712. [[CrossRef](#)]
17. Saka, A.B.; Chan, D.W. Profound barriers to building information modelling (BIM) adoption in construction small and medium-sized enterprises (SMEs) An interpretive structural modelling approach. In *Construction Innovation*; Emerald Group Publishing Limited: Leeds, UK, 2020; Volume 20, pp. 261–284.
18. Dannoun, Y. Application of supply chains management in construction project: A review in the compatibility between the procurements and implementation process. *Int. J. Adv. Eng. Sci. Appl.* **2022**, *3*, 18–21. [[CrossRef](#)]
19. Bouabdallaoui, Y.; Lafhaj, Z.; Yim, P.; Ducoulombier, L.; Bennadji, B. Predictive Maintenance in Building Facilities: A Machine Learning-Based Approach. *Sensors* **2021**, *21*, 1044. [[CrossRef](#)] [[PubMed](#)]
20. Chew, K.C. Singapore's strategies towards sustainable construction. *IES J. Part A Civ. Struct. Eng.* **2010**, *3*, 196–202. [[CrossRef](#)]
21. Ekin, S. Prompt engineering for ChatGPT: A quick guide to techniques, tips, and best practices. *TechRxiv* **2023**. Available online: <https://www.techrxiv.org/doi/full/10.36227/techrxiv.22683919.v2> (accessed on 1 June 2024).
22. Mesko, B. Prompt engineering as an important emerging skill for medical professionals: Tutorial. *J. Med. Internet Res.* **2023**, *25*, e50638. [[CrossRef](#)] [[PubMed](#)]
23. OpenAI. Introducing ChatGPT. 30 November 2022. Available online: <https://www.openai.com/research/introducing-chatgpt> (accessed on 1 June 2024).
24. Heston, T.F.; Khun, C. Prompt Engineering in Medical Education. *Int. Med. Educ.* **2023**, *2*, 198–205. [[CrossRef](#)]
25. Ortolan, P. Optimizing Prompt Engineering for Improved Generative AI Content. Master's Thesis, Universidad Pontificia Comillas, Madrid, Spain, 2023. Available online: <http://hdl.handle.net/11531/80629> (accessed on 1 June 2024).
26. OpenAI. Prompt Engineering. 2023. Available online: <https://platform.openai.com/docs/guides/prompt-engineering> (accessed on 1 June 2024).
27. B4UBUILD. A Sample Residential Construction Schedule. 2024. Available online: <https://www.b4ubuild.com/resources> (accessed on 1 June 2024).
28. Dipankar, P. Multipurpose Building Construction Plan (Schedule/Task List). 2020. Available online: <https://www.researchgate.net> (accessed on 1 June 2024).
29. NJ Department of Transportation Headquarters Complex. SOP Main Office Building and Engineering and Operations Lobby Reconfiguration. 2020. Available online: <https://nj.gov> (accessed on 1 June 2024).
30. Seppänen, O.; Aalto, E. A case study of line-of-balance based schedule planning and control system. In Proceedings of the 13th Annual Conference of the International Group for Lean Construction, Sydney, Australia, 19–21 July 2005.
31. Jarabulus, M. Application of Monte Carlo simulation and PERT/CPM techniques in planning of construction projects: A Case Study. *Period. Eng. Nat. Sci.* **2017**, *5*. [[CrossRef](#)]

32. Occupational Safety and Health Administration (OSHA). Prevention Videos (v-Tools), Construction Hazards. 2024. Available online: <https://www.osha.gov> (accessed on 1 June 2024).
33. Occupational Safety and Health Administration (OSHA). Hazard Communication. (29 C.F.R. § 1910.1200). U.S. Department of Labor. Available online: <https://www.osha.gov/laws-regs/regulations/standardnumber/1910/1910.1200> (accessed on 1 June 2024).

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.