Alexandria Engineering Journal 116 (2025) 672–698

Contents lists available at ScienceDirect

Alexandria Engineering Journal

journal homepage: www.elsevier.com/locate/aej

Original article

Generative artificial intelligence in construction: A Delphi approach, framework, and case study

Ridwan Taiwo [a,b], Idris Temitope Bello [c], Sulemana Fatoama Abdulai [a], Abdul-Mugis Yussif [a], Babatunde Abiodun Salami [d], Abdullahi Saka [e], Mohamed El Amine Ben Seghier [f,*], Tarek Zayed [a]

[a] Department of Building and Real Estate, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong
[b] Institute of Construction and Infrastructure Management, ETH Zurich, Stefano-Franscini-Platz 5, Zurich, Switzerland
[c] Aerospace and Mechanical Engineering, University of Oklahoma, Norman, OK 73019, USA
[d] Cardiff School of Management, Cardiff Metropolitan University, Llandaff Campus, Cardiff CF5 2YB, United Kingdom
[e] Westminster Business School, University of Westminster, London, United Kingdom
[f] Department of Built Environment, Oslo Metropolitan (OsloMet) University, Oslo, Norway

ARTICLE INFO

Keywords:
Generative artificial intelligence
Generative pre-trained transformer
Large language model
Multimodal AI
Retrieval augmented generation
Construction industry
GenAI
RAG
LLM
GPT
ChatGPT

ABSTRACT

The construction industry plays a crucial role in the global economy, contributing approximately $10 trillion and employing over 220 million workers worldwide, but encounters numerous productivity challenges with only 1 % annual growth compared to 2.8 % for the global economy. These challenges span various processes, including design, planning, procurement, inspection, and maintenance. Generative artificial intelligence (GenAI), capable of producing new and realistic data or content such as text, images, videos, or code from given inputs or existing knowledge, presents innovative solutions to these challenges. While there is an increasing interest in the applications of GenAI in construction, a detailed analysis of its practical uses, advantages, and areas ripe for development is still evolving. This study contributes to this emerging area by offering an insightful analysis of the current state of generative AI in construction. It has three objectives: (1) to identify and categorize the existing and emerging generative AI opportunities and challenges in the construction industry via a Delphi study; (2) to propose a framework enabling construction firms to build customized GenAI solutions; and (3) to illustrate this framework through a case study that employs GenAI model for querying contract documents. Through systematic review and expert consultation, the study identified 76 potential GenAI applications across construction phases and 18 key challenges distributed across domain-specific, technological, adoption, and ethical categories. The case study's findings show that retrieval augmented generation (RAG) improves the baseline large language model (LLM), GPT-4, by 5.2, 9.4, and 4.8 % in terms of quality, relevance, and reproducibility. The study recommends a structured approach to GenAI implementation, emphasizing the need for domain-specific customization, robust validation protocols, and careful consideration of ethical implications. This study equips academics and construction professionals with a comprehensive analysis and practical framework, facilitating the integration of GenAI techniques to enhance productivity, quality, safety, and sustainability across the construction industry.

1. Introduction

The construction industry is a cornerstone of the global economy, contributing approximately $10 trillion or 13 % to the global GDP in 2019 [1–3]. It is also a major employer, providing jobs to over 220 million workers worldwide and acting as a key driver of employment [4]. This industry is composed of various sub-sectors, including infrastructure engineering, residential, commercial, and industrial building construction [1–3]. Building construction stands out as the largest and most varied sub-sector, representing around 40 % of the global construction output and half of the construction employment globally [5].

Building construction is recognized for its complexity and dynamism, involving a wide array of stakeholders such as owners, architects,

* Corresponding author.
E-mail address: moseg7662@oslomet.no (M.E.A. Ben Seghier).

https://doi.org/10.1016/j.aej.2024.12.079

Available online 9 January 2025

engineers, contractors, subcontractors, suppliers, and regulators[6,7]. The construction process demands coordinating and integrating numerous activities, including design, planning, scheduling, procurement, fabrication, installation, inspection, and maintenance [8,9]. Additionally, it involves the generation and utilization of a vast amount of data, ranging from drawings and specifications to contracts, reports, invoices, and photographs[10–12]. The quality, efficiency, and sustainability of building construction are heavily dependent on the effective management and use of these activities and data. Despite its significance, the construction industry encounters numerous challenges that impede its performance and productivity. These challenges are prevalent across all lifecycle stages of construction, such as design, construction, procurement and supply chain, fabrication and installation, and inspection and maintenance, as illustrated in Fig. 1 [13–15]. Each stage is characterized by complex and dynamic processes that necessitate the coordination and integration of diverse resources, disciplines, and stakeholders while also managing a variety of constraints, uncertainties, and changes. These challenges present considerable difficulties and risks to the construction industry, leading to reduced productivity, increased costs, prolonged delays, diminished quality, and greater environmental impact. According to a report by McKinsey, the global construction industry has an average annual productivity growth of only 1 %, compared to 2.8 % for the total world economy and 3.6 % for manufacturing. The report also estimates that the global construction industry could save up to $1.6 trillion per year by improving its productivity to the level of other sectors [1].

Addressing these challenges and improving the performance and productivity of the building construction industry requires innovative and disruptive solutions that can leverage the power of data and technology. One of the most promising and emerging solutions is generative artificial intelligence (AI) [16–18]. Generative AI (GenAI) is a branch of AI that aims to create novel and realistic data or content, such as text, image, video, audio, or code, based on some input or prior knowledge [19–21]. GenAI can be seen as the opposite of discriminative AI, which aims to classify or recognize data or content, such as identifying objects in an image or translating text from one language to another. More elaboration on the foundational concepts of GenAI relevant to this discourse is provided in the supplementary material.

GenAI encompasses large language models (LLMs) and multimodal GenAI models (multimodal GenAI), both of which are neural network models trained on extensive datasets. LLMs, such as those trained on texts from books, articles, websites, and social media, are adept at capturing natural language's semantic and syntactic nuances [22–24]. Multimodal GenAI, on the other hand, is an advanced algorithm that can generate not only text but also images, audio, video, or code. They learn general representations from large corpora of varied content during pretraining, which allows them to produce original outputs across different modalities. GenAI models are built upon the foundational algorithms discussed in the supplementary material of this study and include examples like GPT-4 and Codex, which generate coherent text

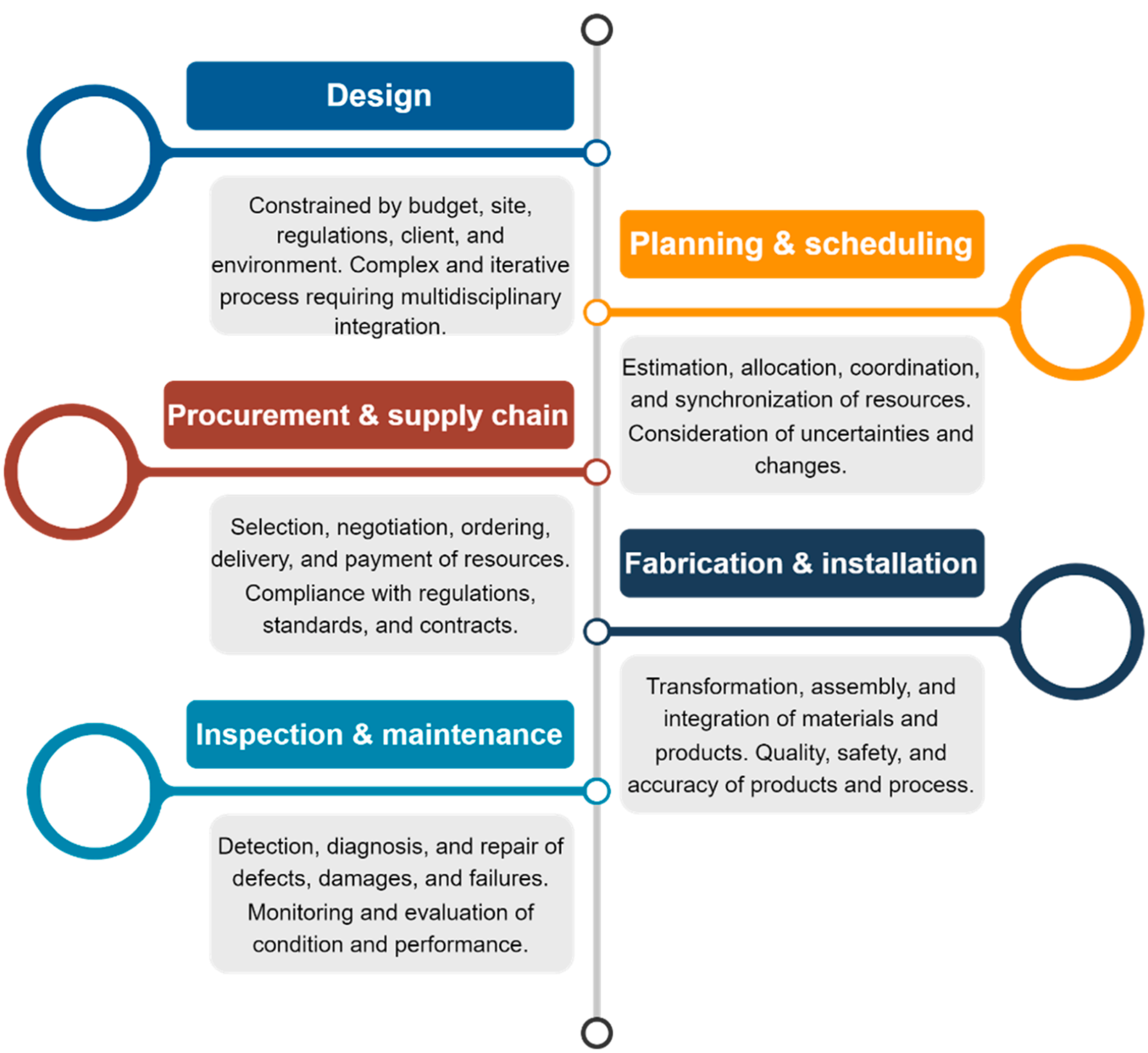


Fig. 1. Major activities and challenges encountered at each phase of a building construction cycle.

and functional code, respectively [25,26]. Similarly, DALL-E 3 and Imagen can create photorealistic images from text descriptions [27,28], while models trained on audio can synthesize natural human speech or music. Multimodal GenAI models' versatility extends to multimodal generation, such as pairing input stories with appropriate images or matching lyrics with suitable music compositions. The hallmark of multimodal GenAI is its colossal scale in terms of model size, computational requirements, and training data volume, which facilitates a wide-ranging and inventive generative capacity that spans text, images, video, code, and integrated multimodal outputs. This positions them as a comprehensive generative toolbox, pre-trained on diverse big data for multifaceted applications. GenAI's adoption across various disciplines demonstrates its versatility: in healthcare, it generates synthetic medical images for privacy and dataset augmentation [29]; in pharmaceuticals, it aids in developing new molecular structures for drug discovery [30,31]; in business, it supports generative design for product ideation [32]; in social sciences, it assists in historical document restoration and synthetic population data generation [33]; and in academia, it contributes to personalized educational tools and interactive simulations [34]. Table 1 provides an overview of recent GenAI models, detailing their developers, training parameters, release years, and accessibility.

Although GenAI applications are still at an infant stage in the construction industry, a few studies exist on their usage for construction-related work [25,58]. As such, [59] reviewed applications of GenA I for developing and enhancing structural designs and how they could help improve accuracy in the design process. [26] presented an overview of the potential applications of Generative Pre-trained Transformer (GPT) models across the lifecycle of a construction project and a case study for material selection. Further, text-based opportunities and a limited number of challenges of adopting GenAI in the construction industry were presented by [60]. These studies have laid important groundwork, but their scope remains limited to specific applications or narrow aspects of GenAI implementation.

Table 1
GenAI models released in recent years.

Models	Developer	Training parameter (Billion)	Release year	Access	Ref.
Llama−3.2	Meta	1, 3, 11, and 90	2024	Open source	[35]
Grok	xAI	314	2024	Open source	[36]
Flux-pro	Black Forest Lab	12	2024	API	[37]
Claude−3.5-Opus	Anthropic	-	2024	API	[38]
Claude−3.5-Haiku	Anthropic	-	2024	API	[38]
Claude−3.5-sonnet	Anthropic	-	2024	API	[38]
o1-mini	OpenAI	-	2024	API	[39]
o1-preview	OpenAI	-	2024	API	[39]
GPT−4o	OpenAI	-	2024	API	[40]
GPT−4	OpenAI	-	2023	API	[41]
Gemini Pro	Google DeepMind	-	2023	Open source	[42]
Llama 2	Meta	7, 13, 70	2023	Open source	[43]
PaLM	Google	540	2022	Open source	[44]
DALLE−3	OpenAI	-	2023	API	[27]
SDXL	Stability AI	2.6	2023	Open source	[45]
DALLE−2	OpenAI	3.5	2022	API	[46]
Dreamfusion	Google	-	2022	API	[47]
Flamingo	Google DeepMind	80	2022	API	[48]
Phenaki	Google	1.8	2022	API	[49]
Codex	OpenAI	12	2021	API	[50]
Galactica	Meta	120	2022	Open source	[51]
AudioLM	Google	0.6	2022	API	[52]
DALL-E	OpenAI	12	2021	API	[53]
BART	Facebook AI	0.4	2019	Open source	[54]
T5	Google	11	2019	Open source	[55]
GPT−3.5	OpenAI	175	2022	API	[56]
GPT−2	OpenAI	1.5	2019	Open source	[57]
XLNet	Google	0.34	2019	Open source	[52]

Despite the potential and promise of GenAI for the construction industry, more systematic and comprehensive literature needs to be compiled that identifies and analyses the current state, opportunities, and challenges of GenAI in this domain. Most existing literature focuses on specific applications or aspects of GenAI (such as GPT models or text-text models) without considering the broader and holistic picture of GenAI in the construction industry. Moreover, there is an urgent need for practical and actionable guidance on implementing and deploying GenAI solutions in the construction industry. This is particularly crucial for construction firms that may lack sufficient data, expertise, or resources to develop their own GenAI models from scratch. The industry requires a framework that bridges the gap between theoretical possibilities and practical implementation, addressing everything from data collection and preprocessing to model deployment.

Therefore, this study aims to provide a holistic analysis of GenAI in the construction industry, with the following objectives:

- To provide opportunities and challenges of applying GenAI in the construction industry.
- To propose a framework enabling construction firms to build customized GenAI solutions using their data by describing and explaining the key steps and components of the framework.
- To demonstrate the proposed framework via a practical use case of developing a tailored GenAI for contract documents.

This study's unique contribution lies in its comprehensive approach that combines theoretical analysis with practical implementation guidance, supported by expert insights and validated through an original case study. Unlike previous works focusing on specific applications or theoretical possibilities, this research provides a complete roadmap for construction firms to successfully implement GenAI solutions while addressing industry-specific challenges and requirements.

The rest of the paper is organized as follows: Section 2 describes the study's methodology, which is explained in four phases. Section 3 presents the literature review results and expert discussion, including the current applications, opportunities, and challenges of GenAI in the construction industry. Section 4 proposes the framework for building custom GenAI in the construction industry and explains the main steps and components of the framework. Section 5 demonstrates the framework via a case study of developing a GenAI model for contract documents and shows the results and outcomes of the case study. Section 6 concludes the paper and provides some directions for future research.

2. Methodology

A four-phase approach is adopted to achieve the objectives of this study. Fig. 2 visualizes these phases, including systematic literature review and retrieval, expert discussion and review, a framework for developing a custom GenAI model in the construction industry, and a case study.

I. **Phase 1 – Systematic literature retrieval and review**: The first step in this phase involves selecting appropriate databases for the literature search. Scopus, Web of Science, and ScienceDirect were chosen due to their broad coverage and rigorous indexing of peer-reviewed publications [61,62]. Keyword identification was

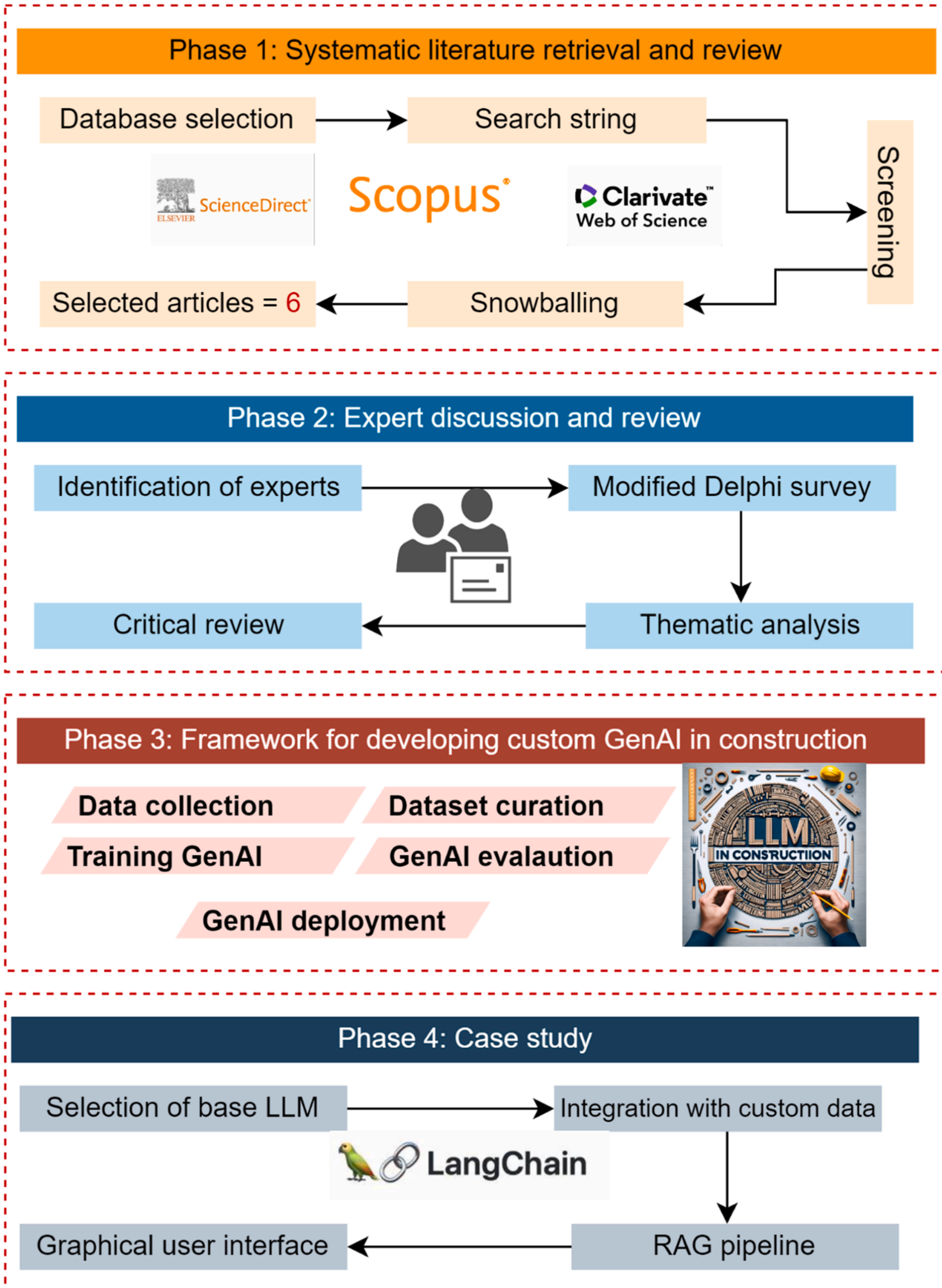


Fig. 2. Research framework.

conducted iteratively to capture relevant studies at the intersection of GenAI and the construction industry. The final search string consisted of ["Construction industry" OR "architecture engineering and construction industry" OR "AEC industry" OR "AECO industry"] AND ["Generative AI" OR "GenAI" OR "GENAI" OR "Bard" OR "Gemini" OR "GPT" OR "GPT-1" OR "GPT-2" OR "GPT-3" OR "GPT-3.5" OR "GPT-4" "ChatGPT" OR "Transformer" OR "GPT-4" OR "Llama" OR "LamDA"]. This search returned 79 initial results. The search results were narrowed to 10 potentially relevant studies based on the title and abstract screening. An in-depth review found that only four (4) were original research articles, with two review articles, and the full text of the rest was unavailable or written in languages other than English. Snowball searching expanded the final pool to six original research articles at the intersection of GenAI and construction.

II. **Phase 2 – Expert discussion and review**: The limited literature identified in Phase 1 highlighted the need to supplement with expert perspectives, given the nascent state of GenAI adoption in construction. To elicit diverse insights, 15 experts with backgrounds spanning AI research and construction industry practice were identified. Invitations were sent to participate in the study, with 11 experts accepting for a 73 % response rate. This panel encompassed university professors in AI and construction engineering, technology directors from major construction firms, and founders of AI startups targeting the architecture, engineering, and construction (AEC) industry. Table 2 shows the demographic details of the experts. A Delphi survey consisting of three rounds was conducted with the panel to identify opportunities and challenges of applying GenAI in the construction industry. Thematic analysis was then used to extract common themes from the qualitative responses. This involved codifying the experts'

Table 2
Demographics of the experts.

Category	Profile	Frequency	Percentage (%)
Sector	Academia	6	54
	AI industry	5	46
Experience	5–10 years	2	18
	11–15 years	3	27
	16–20 years	4	36
	20 and above	2	18
Highest degree	Bachelor	1	9
	Master	3	28
	Doctorate	7	63
Background	Architecture	2	18
	Civil engineering	3	27
	Computer science	3	27
	Construction management	3	27

opinions and aggregating them into categories through an iterative process. The goal was to determine areas of consensus as well as unique perspectives.

III. **Phase 3 – Framework for developing custom GenAI in the construction industry**: This phase involved synthesizing the literature and expert findings into a methodology construction firms can follow to build custom GenAI solutions using their proprietary data. The framework encompasses construction data collection, dataset preprocessing, training the custom GenAI model, evaluation, and deployment steps. The details of the framework are presented in Section 4.

IV. **Phase 4 – Case study**: A case study was conducted using GenAI to query contract documents and demonstrate practical application. The first step involved the selection of the base LLM architecture. OpenAI's GPT-4 model was chosen as the base model due to its state-of-the-art natural language generation capabilities. A retrieval-augmented generation (RAG) system was implemented to improve the base LLM further. This mitigated hallucinated text by grounding outputs in relevant dataset examples. LangChain Library was employed for the development [63]. The performance of the customized LLM was evaluated, and a graphical user interface was developed using Streamlit [64]. This interactive web application enabled testing of the customized GenAI model through prompts.

3. Results

3.1. Current applications

This section presents the application of GenAI in the form of LLMs in the construction industry. Based on the systematic review, only six peer-reviewed articles exploring the uses of LLMs in construction were identified. No articles relating to other GenAI models, such as large image and video models, were found. The six articles are summarized in Table 3, including their objective, methods, and contributions. The reviewed studies demonstrate emerging applications of LLMs, such as GPT-based models for construction tasks, including virtual assistance, sequence planning, schedule generation, hazard recognition, risk assessment, and project planning [65–67]. The contributions highlight the potential for LLMs to enhance productivity, accuracy, and automation in areas like information retrieval, education/training, and documentation review. However, the limited number of studies indicates that the adoption of modern GenAI in construction is still in the very early stages. Significant research is needed to develop customized LLMs and multimodal GenAI models for the industry and validate their capabilities in dealing with real-world problems.

3.2. Opportunities

The discussions conducted with experts revealed 76 potential opportunities to deploy GenAI across construction, categorized by input-output capabilities. Sections 3.2.1 through 3.2.9 extensively examine applications of generative texts, images, and videos. While generative models can also synthesize audio, experts advised that video generation can serve dual visual and auditory content purposes as construction relies heavily on visual data like drawings, photos, animations, and written and verbal communications, generative modes spanning text, images, and video were seen as most directly relevant. Despite significant progress, there are still limitations, particularly when generating complex images and videos. Table 4 summarizes leading generative

Table 3
Summary of current applications of LLM in the construction industry.

Ref.	Objective	Methods	Contributions
[58]	Development of a dynamic prompt-based virtual assistant framework for BIM information search	The framework integrates BIM and GPT technologies for an NL-based interface. Dynamic prompt-based process interprets NL queries, retrieves information, and delivers responses.	The framework's application improves information search speed, accuracy, and user experience.
[68]	Development of RobotGPT for automated sequence planning in robotic assembly for construction tasks.	RoboGPT is a system that uses ChatGPT for automated sequence planning in robot-based construction assembly.The experimental evaluation included two case studies and 80 trials involving real construction tasks.	RoboGPT-driven robots can handle complex construction operations and adapt to changes on the fly.
[25]	Generation of a construction schedule for a project	ChatGPT is employed to generate a construction schedule for a simple project.A survey was conducted to evaluate output quality and participants' experience.Parameters used to evaluate results include accuracy, efficiency, clarity, coherence, reliability, relevance, consistency, scalability, and adaptability.	The use of LLM to enhance construction schedules workflow.
[69]	The use of ChatGPT for improving hazard recognition on construction site	The investigation involved 42 students in a construction program.Pre- and post-intervention hazard recognition abilities were measured.	The potential of employing ChatGPT for safety education and training.
[70]	Automated classification of contractual risk clauses	The BERT method is used for clause classification in construction specifications.Seven risk categories were identified: payment, temporal, procedure, safety, role and responsibility, definition, and reference.	The model improves the construction specification review process and risk management.
[71]	Automatic matching of look-ahead planning tasks to master scheduled activities	Both location-based and distance-based matching followed were employed.GPT−2 was used for final matching.	Auto-alignment of long-term and short-term plans in construction projects

Table 4
GenAI models for various input-output types.

Input-output type	Model	Developer	Ref.
Text to text	Claude−3.5, GPT−4o, Gemini Pro	Anthropic, OpenAI, Google's DeepMind	[72, 73]
Text to image	DALL-E 3, Flux	OpenAI, Black Forest Lab	[27]
Text to video	CogVideo, Lumiere	Nightmareai, Google Research	[74, 75]
Image to text	GPT−4, Gemini Pro	OpenAI, Google's DeepMind	[72, 73]
Image to image	Pix2Pix	Berkeley AI Research	[77]
Image to video	Stable Video Diffusion	Stability AI	[78]
Video to text	VideoCoCa	Google Research	[79]
Video to image	-	-	
Video to video	Lumiere, Gen−2	Google Research, Runway	[75, 76]

models for different data types. LLMs like GPT-4o and Gemini Pro demonstrate proficiency in text synthesis [72,73]. DALL-E 3 produces images from text captions [27]. Video generation models like CogVideo, Lumiere, and Stable Diffusion show promise but are still being refined [74,75]. Although there are shortcomings, the pace of progress makes GenAI a promising technology for transforming the construction industry. If trained on sufficient domain data, text generation achieves high coherence and accuracy. Photorealistic image synthesis provides value in design and documentation use cases. Video capabilities lag but rapidly improve through advances like higher resolution GANs [75,76].

3.2.1. Text to text

GenAI revolutionizes the construction industry by converting textual data into advanced textual outputs, assisting in many tasks in the construction project phases. Table 5 provides a detailed overview of the potential applications of text generation in the construction industry, categorized using the different project phases. In pre-construction, it can help generate feasibility study summaries, ensure regulatory compliance, and automate proposal/bid drafting [80]. Drafting daily progress reports, specifications, task instructions, and other documents can be automated during construction (see Fig. 3). Post-construction opportunities include creating inspection reports, punch lists, operation and maintenance manuals, reviewing warranty/compliance letters, and translating documents. Other cross-cutting text applications are information retrieval through natural language queries and translation into multiple languages [17]. With proper training in technical corpora, they can translate industry insights directly into clear, accurate documents without tedious hands-on work. Realizing this potential requires careful, prompt engineering and alignment with construction linguistic patterns and technical jargon.

3.2.2. Text to image

AI's text-to-image conversion provides innovative possibilities in the construction field, including the ability to visualize pre-construction architectural ideas, assist in making real-time construction choices, and enhance marketing materials once construction is completed. The potential opportunities for generating images via text prompting are shown in Table 6. Generating images from text has broad applicability in construction projects. Pre-construction applications include creating visualizations from site descriptions for selection and planning [81]. Text prompts can also render architectural concepts and project models (Fig. 4). During construction, progress visualization, equipment layouts, and safety illustrations can be automated from textual inputs. Fig. 5 shows a visualization of construction progress through different stages of execution. Post-construction use cases involve as-built visualization, usage guidelines, and renovation proposals. With appropriate training,

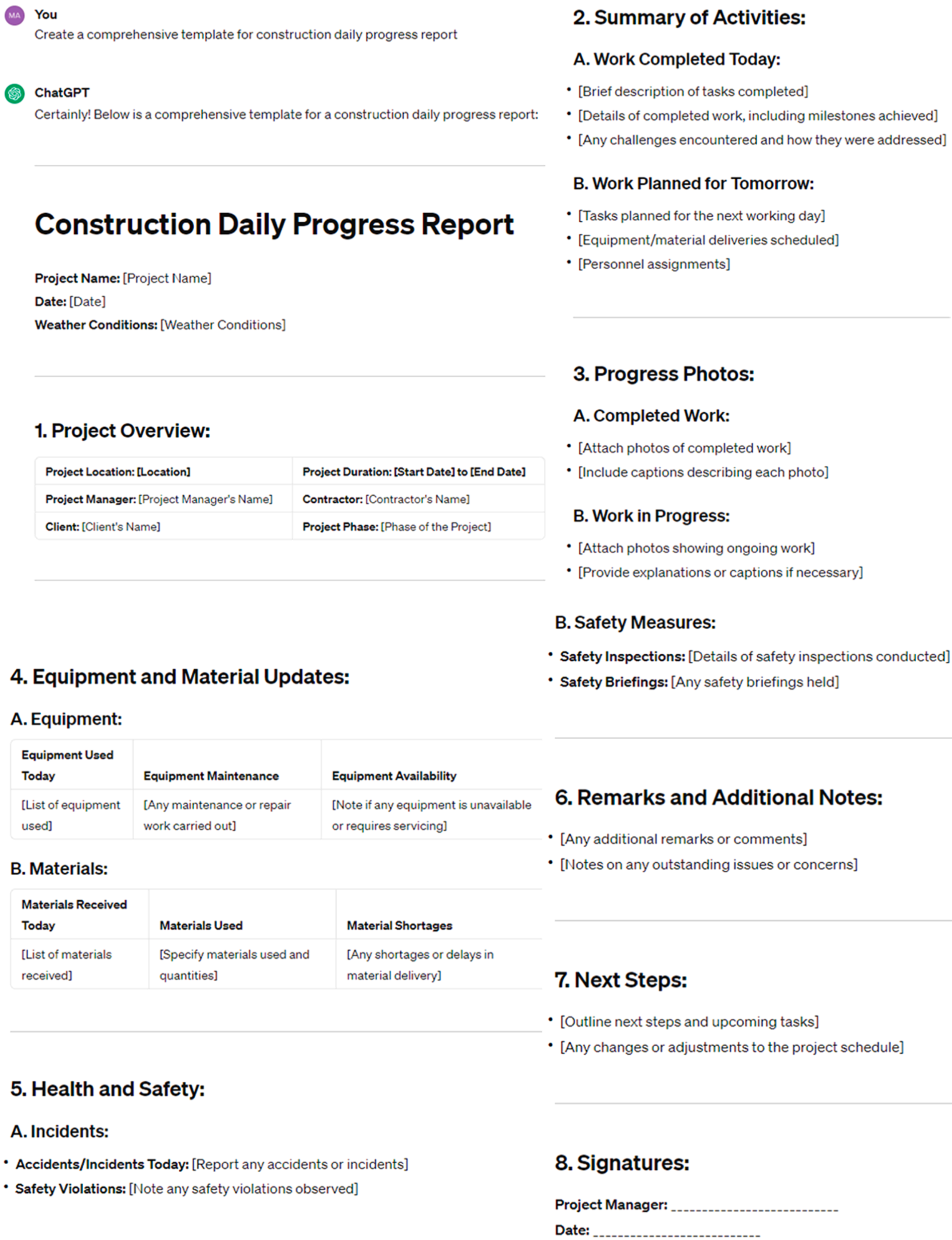
Table 5
Potential GenAI opportunities in the construction industry for text-text model type.

Potential opportunity	Description	Project phase
Generation of the feasibility report summary	Summarize extensive feasibility reports and extract key insights and recommendations for informed decision-making during the project initiation.	Pre-construction
Documentation of regulatory compliance	Leverage GenAI to assist in creating documents that ensure compliance with regulatory requirements, a crucial task in the pre-construction planning phase.	Pre-construction
Preparation of proposal/bid	Apply GenAI to assist in the preparation of proposals and bids by automatically generating well-structured and persuasive text content.	Pre-construction
Generation of daily progress report	Create a daily progress report template summarizing on-site activities and achievements during construction.	Construction
Refinement of construction specifications	Utilize GenAI to refine and enhance construction specifications, ensuring clarity and accuracy in the documentation of materials, methods, and standards.	Construction
Task Assignment and Communication	Facilitate task assignment and communication by automatically generating clear and detailed instructions for construction teams through GenAI.	Construction
Summarization of as-built documents	Summarize the extensive as-built documentation, providing a condensed overview of the final constructed project for post-construction analysis.	Post-construction
Generation of facility maintenance manual	Automate the generation of comprehensive facility maintenance manuals based on the final as-built documentation.	Post-construction
Review of warranty and compliance document	Utilize GenAI to review warranty and compliance documents, summarizing critical information and ensuring adherence to post-construction requirements.	Post-construction
Language translation and localization	Translate text content between different languages, aiding in global collaboration and communication.	All
Information retrieval and knowledge discovery	Enhance contextual search capabilities by using GenAI to understand and respond to natural language queries, improving the accuracy of information retrieval.	All

models like DALL-E can translate construction domain language into detailed visuals through well-prompted texts. This technology allows people without expertise to readily obtain visual depictions by articulating what they wish to see in plain language. Automating this linkage between vision and language can make project information more accessible while freeing worker time [26].

3.2.3. Text to video

The utilization of GenAI to transform textual information into dynamic video content offers numerous benefits. Table 7 summarizes key opportunities for text-to-video generation in construction based on the expert discussion. During the pre-construction phase, introductory site exploration videos and animated project concept videos could be synthesized from text to aid scope planning and stakeholder intelligence. During construction, step-by-step equipment operation tutorials and safety training animations could be generated from manuals and textual hazard narrations, respectively [82]. Progress update videos compiled from schedules and logs would help keep stakeholders informed with



You

Create a comprehensive template for construction daily progress report

ChatGPT

Certainly! Below is a comprehensive template for a construction daily progress report:

Construction Daily Progress Report

Project Name: [Project Name]
Date: [Date]
Weather Conditions: [Weather Conditions]

1. Project Overview:

Project Location: [Location]	**Project Duration: [Start Date] to [End Date]**
Project Manager: [Project Manager's Name]	**Contractor:** [Contractor's Name]
Client: [Client's Name]	**Project Phase:** [Phase of the Project]

2. Summary of Activities:

A. Work Completed Today:

- [Brief description of tasks completed]
- [Details of completed work, including milestones achieved]
- [Any challenges encountered and how they were addressed]

B. Work Planned for Tomorrow:

- [Tasks planned for the next working day]
- [Equipment/material deliveries scheduled]
- [Personnel assignments]

3. Progress Photos:

A. Completed Work:

- [Attach photos of completed work]
- [Include captions describing each photo]

B. Work in Progress:

- [Attach photos showing ongoing work]
- [Provide explanations or captions if necessary]

4. Equipment and Material Updates:

A. Equipment:

Equipment Used Today	Equipment Maintenance	Equipment Availability
[List of equipment used]	[Any maintenance or repair work carried out]	[Note if any equipment is unavailable or requires servicing]

B. Materials:

Materials Received Today	Materials Used	Material Shortages
[List of materials received]	[Specify materials used and quantities]	[Any shortages or delays in material delivery]

5. Health and Safety:

A. Incidents:

- **Accidents/Incidents Today:** [Report any accidents or incidents]
- **Safety Violations:** [Note any safety violations observed]

B. Safety Measures:

- **Safety Inspections:** [Details of safety inspections conducted]
- **Safety Briefings:** [Any safety briefings held]

6. Remarks and Additional Notes:

- [Any additional remarks or comments]
- [Notes on any outstanding issues or concerns]

7. Next Steps:

- [Outline next steps and upcoming tasks]
- [Any changes or adjustments to the project schedule]

8. Signatures:

Project Manager: _________________________
Date: _________________________

Fig. 3. Daily progress report template generated using GPT-4 via ChatGPT interface.

Table 6
Potential GenAI opportunities in the construction industry for text-image model type.

Potential opportunity	Description	Project phase
Site visualization and selection	Create visual representations of potential construction sites based on text descriptions, aiding the decision-making process during site selection.	Pre-construction
Architectural concept rendering	Transform textual architectural concepts into visual renderings, providing stakeholders with a clear preview of the proposed designs.	Pre-construction
Interactive project models	Utilize GenAI to convert project descriptions into interactive 3D models, allowing stakeholders to explore and engage with the project before construction begins.	Pre-construction
Construction progress visualization	Implement GenAI to generate visual representations of construction progress based on textual updates, providing stakeholders with a visual timeline of the project.	Construction
Material and equipment layouts	Through generative AI, visual layouts of materials and equipment based on textual descriptions are created, optimizing their placement on the construction site.	Construction
Safety procedure illustrations	Apply GenAI to convert text-based safety procedures into visual illustrations, enhancing comprehension and adherence to safety protocols on the construction site.	Construction
As-Built visualization	Transform the as-built documentation into visual representations, aiding in the visualization and analysis of the final construction.	Post-construction
Facility usage guidelines	Create visual guidelines for facility usage based on textual documentation, ensuring clear communication of post-construction guidelines.	Post-construction
Renovation proposal visualizations	Generate visual representations of proposed renovations, aiding decision-making during the post-construction phase.	Post-construction
Project timeline infographics	Convert textual project timelines into visual infographics, providing an easily understandable overview for all project phases.	All
Project dashboard visuals	Generate visual representations for project dashboards based on textual data, offering stakeholders an intuitive and informative overview of project metrics.	All

matching visual updates. Post-construction use cases include creating instructional facility usage videos from the documentation. With appropriate training data, text-to-video models can translate construction domain language into vivid animations and live footage [74,75]. Rather than relying solely on static diagrams and dense text, bringing instructions and processes to life through AI-generated videos makes project information more engaging. Dynamic video tutorials personalized via text to each situation may enhance comprehension and learning for safety training and equipment operation. Automating the linkage between textual descriptions and video footage also frees workers' time spent manually storyboarding and editing visualizations. As text-to-video generation techniques continue advancing in resolution and realism, the applications across the construction project lifecycle will expand.

3.2.4. Image to text

Converting images into text descriptions has valuable applications throughout construction project phases. Table 8 summarizes potential use cases validated by the expert discussion. Pre-construction opportunities include extracting measurements, boundaries, and other information from land surveys and blueprints. During construction, daily site photos could be analyzed to generate progress reports. Images of materials and equipment could develop real-time quality and inventory assessments via GenAI[83]. Fig. 6 displays an image description generated by Gemini Pro as part of a daily visual report. Upon close inspection, the model accurately captures fine-grained details in the image, including identifying the specific brand and model of the construction equipment. This demonstrates Gemini Pro's capability to produce descriptive text summarizing critical visual information [73]. Post-construction applications involve extracting as-built details from archival photos and making warranty documentation from damaged images. Cross-cutting use cases include automating visual inspection reports across phases. With proper training, image captioning techniques can translate construction graphics and photos into structured textual information. This eliminates tedious manual efforts to log and convey visual observations. Models such as GPT-4o can analyze everyday images and accurately describe prominent objects, actions, and scenery [72].

3.2.5. Image to image

The image-to-image capabilities of GenAI are crucial in construction as they facilitate design adjustments, on-site issue solutions, and the visualization of future upgrades. Table 9 details potential applications of image-image models in the construction industry. Pre-construction use cases include adapting architectural visualizations into different desired art styles and refining scanned maps into clear site plans. During construction, input architectural plans and sketches could be auto-modified to match ongoing changes on-site [59]. Material texture libraries could help generate realistic composite renderings from sample images. Post-construction applications involve visualizing the restored building state from damage assessment images and landscape enhancements. Improving image quality and resolution are potential applications across all phases of construction projects. Techniques such as pix2pix GANs demonstrate capabilities to transform input images while preserving essential content structure [77]. By learning alignments between construction image domains during training, models can translate inputs into desired stylistic, structural, or conceptual outputs. This allows the adaptation of visual data into appropriate formats for downstream usage, reducing repetitive manual editing. For instance, rough sketches produced during early design phases can be refined into polished architectural visualizations or engineering schematics. Images captured on-site can be adapted to match design intent, even when physical conditions vary. Continued advances in high-resolution GANs will further expand the potential for image-to-image synthesis to enhance visual media throughout construction projects.

3.2.6. Image to video

Converting static images into dynamic videos opens up impactful possibilities across the construction project lifecycle. Table 10 summarizes potential applications in this area. During pre-construction, static architectural concept images could be converted into engaging animated walkthroughs and fly-throughs to showcase designs. Aerial site photos could also produce simulated planning and development timelapses [84]. Safety incidents could be recreated on active construction sites based on analysis of images of unsafe conditions to improve hazard awareness through vivid video representations. Time-lapse build videos compiled from daily construction photos help visually track project progression. Post-construction use cases include generating promotional experience videos from facility images and collecting recap documentary videos from archival visuals. State-of-the-art generative video models demonstrate increasing capabilities to animate photo-realistic footage from sparse image inputs [78]. AI systems can extend single images into complete video sequences with convincing continuity and realism by learning to extrapolate motion and physical interactions [85]. Construction visuals contain extensive intrinsic structures that video generation models can leverage to produce

Fig. 4. Image of 3D BIM Model Generated Using DALL-E 3.

meaningful video representations without full frame-by-frame supervision. Converting images into dynamic videos helps improve engagement and understanding compared to static depictions alone. As the coherence and resolution of image-to-video models continue improving, their potential applications in construction for bringing visuals to life will grow.

3.2.7. Video to text

AI transcription of video-to-text revolutionizes the construction industry by providing comprehensive documentation throughout the construction process. Potential applications of video-to-text conversion are shown in Table 11. During pre-construction, generative models could auto-transcribe kickoff meetings and regulatory compliance tutorial videos into concise text records. Safety briefing videos and daily progress meeting discussions on active construction sites could be translated into text summaries for distribution to wider stakeholders [82]. Post-construction commissioning and inspection videos also contain valuable verbal feedback that video-to-text techniques can structure into reports. Across phases, comprehensively transcribing archived project videos into indexed, searchable documentation enables robust retrospective analysis [83]. Models such as VideoCoCa can transcribe technical construction multimedia while filtering out irrelevant background noise [79]. The text outputs synthesize the key details and language from videos without needing to review hours of footage. This allows scaling-extraction of vital audio information in rich multimedia that construction teams continuously generate. When deployed with proper data controls, video-to-text AI can unlock new levels of value from archived construction data without demanding extensive manual effort.

3.2.8. Video to image

As shown in Table 12, extracting key representative images from construction videos offers value across construction projects. Exploring site videos could be condensed into salient snapshots during pre-construction to accelerate assessments. Video conferences discussing design concepts can be automatically packed into a collage of snapshot visuals [86]. In the construction phase, delivery footage could be processed into consolidated photo logs of materials arriving on-site. Aerial construction video can generate periodic bird's-eye progress views [84]. Post-construction applications include extracting instructional stills from facility tutorials. Compiling time-lapse visual collages from archival videos can summarize entire project journeys. Video summarization techniques such as recurrent auto-encoders demonstrate capabilities to identify important frames that distill key visual concepts from longer video sequences [87]. Through the application of these techniques to construction footage, essential moments can be extracted without the necessity for manual video scrubbing. The representative thumbnail images could support rapid video review and summarization. They also integrate more seamlessly into reports and presentations than video embeds. Further innovation in dense video understanding and summarization will continue expanding the capabilities for automating the extraction of impactful visuals from construction multimedia.

3.2.9. Video to video

While static outputs enable analysis, video can engage stakeholders through dynamic visualization. Constructing the future requires

Fig. 5. Construction progress visualization generated by DALL-E3.

envisioning it in motion. Generative video-to-video models can help bring these visions to life across the project lifecycle, as depicted in Table 13. During pre-construction planning and bidding, generative models could synthesize simulated construction sequences from source videos to allow interactive visualization of various work strategies and schedules for optimization [88]. On construction sites, input training videos could be adapted into multi-lingual versions translated across diverse crews to increase accessibility and comprehension. Architectural visualization videos, as-built documentation, and sensor data could be synthesized into lifecycle simulations and digital twin representations to support operations and maintenance. Opportunities exist for efficient review across all phases, such as video quality enhancement, accelerated time-lapses, and video summarization.

By learning spatial-temporal relationships from construction footage, models can extend source videos into modified outputs adapted for downstream requirements. This allows for tailoring visual media for specific applications ranging from training to monitoring to forecasting. As video generation techniques continue advancing, the potential for AI-assisted video remixing and synthesis to enhance multimedia value across the construction project lifecycle will grow substantially [75]. Processing datasets accumulating from the proliferation of construction cameras and sensors using generative video models promises to unlock new visual insights and perspectives.

3.3. Challenges

The adoption of GenAI in the construction industry is growing; thus, it comes with challenges. As presented by the experts during the Delphi

Table 7
Potential GenAI opportunities in the construction industry for text-video model type.

Potential opportunity	Description	Project phase
Site introduction videos	Create introductory videos for potential construction sites, providing stakeholders with visual overviews based on textual descriptions.	Pre-construction
Project concept animation	Transform textual project concepts into animated videos, offering stakeholders a dynamic visualization of the proposed construction.	Pre-construction
Equipment operation guides	GenAI can automatically create step-by-step video tutorials demonstrating equipment use from the text and diagrams in instruction manuals.	Construction
Safety procedure animations	Safety managers could compose comprehensive narrations of hazards and precautions. GenAI can synthesize engaging video footage matching the narration to create safety training materials.	Construction
Progress update videos	Generate progress videos automatically using GenAI using progress reports, schedules, logs, and notes.	Construction
Facility usage instruction videos	Generate instructional videos based on textual documentation for facility usage, ensuring clear communication of post-construction guidelines.	Post Construction
Building update videos	Produce AI-generated videos summarizing facility modifications, upgrades, and status changes over time from text-based building logs for stakeholders.	Post Construction
Project journey montage	Implement GenAI to compile a video montage showcasing the entire project journey, combining textual descriptions and visual elements for a comprehensive overview.	All

Table 8
Potential GenAI opportunities in the construction industry for image-text model type.

Potential opportunity	**Description**	**Project phase**
Land survey data extraction	Analyze land survey images and extract textual data, such as measurements, topographical details, and boundary information.	Pre-construction
3D model specification	Analyze 3D architectural models and automatically generate detailed textual specifications of materials, components, dimensions, etc.	Pre-construction
Blueprint digitization	Automatically convert scanned paper blueprints and hand-drawn sketches into structured digital representations.	Pre-construction
Daily progress image analysis	Analyze daily progress images from construction sites and generate textual reports summarizing the progress, challenges, and achievements.	Construction
Material quality assessment	Examine images of construction materials and generate textual assessments regarding quality, potential issues, and compliance.	Construction
Inventory management	Use AI techniques to automatically catalog on-site equipment, materials, tools, etc., from images and videos into searchable inventory databases.	Construction
As-built documentation text extraction	Analyze images of as-built documentation and extract textual information, facilitating the creation of detailed post-construction reports.	Post-construction
Warranty claim documentation	Explore images of construction components and generate textual documentation for warranty claims, specifying issues and relevant details.	Post-construction
Visual inspection reports	Examine images from visual inspections and generate textual reports, providing detailed information on observed conditions and recommendations.	All

survey, these challenges are multi-faceted, encompassing domain-specific, technological, adoption, and ethical challenges, as shown in Fig. 7. Table 14 provides an overview of these challenges along with practical solutions for addressing them in construction applications. These challenges should be navigated for successful real-world deployment of GenAI in construction. A holistic approach considering technical and non-technical factors is required to overcome barriers and unlock the full potential of LLMs in construction.

3.3.1. Domain-specific challenges

3.3.1.1. Requirement for construction-specific knowledge. Construction is a complex field with intricate technical knowledge required to execute projects safely and efficiently [89,90]. However, most current GenAI techniques rely solely on statistical patterns extracted from data. They need to explicitly encode the nuanced human expertise and domain constraints around structural engineering, materials science, construction codes, aesthetics, machinery, schedules, costs, etc. As a result, generative models trained exclusively on construction data may fail to produce valid, high-quality outputs that align with industry best practices. For example, a generative design model may create a visually appealing 3D building model that violates important structural principles, safety factors, or zoning regulations. The lack of engineering heuristics and constraints leads the unrestrained model to hallucinate flawed plans. Likewise, a generative text model trained only on construction documents will fail to generate specifications or instructions demonstrating a human's comprehension of materials compatibilities, sequences of operations, cost impacts, or equipment capabilities. Generative models need better integration of structured domain knowledge beyond just data patterns to reach their potential in construction. This remains challenging as industry experts' rules and mental models are difficult to codify for machines. Advances in neuro-symbolic AI, modular architectures, and expert-in-the-loop training show promise for imbuing models with more robust construction domain intelligence [91].

3.3.1.2. Handling unstructured and heterogeneous data. Construction data exists in multifaceted, unorganized formats across disparate systems, posing challenges for generative AI. Project information encompasses everything from scanned paper blueprints to 3D BIM models, permits, contracts, change orders, requests for information (RFIs), submittals, specifications, budgets, meeting minutes, multimedia, and more. These data types have different structures, semantics, units, symbols, file formats, and modalities. Generative models like GANs and VAEs struggle to ingest this heterogeneous, unstructured data directly to synthesize coherent outputs. For example, a basic image-to-image model cannot map a 3D BIM model, change order form, and permit application into a unified generated output. The variability across projects also hampers standardized tooling. Each construction firm may have customized data conventions, nomenclatures, templates, and systems tailored to their needs. Creating consolidated datasets from dispersed historical records is arduous. To work around these challenges, purpose-built multi-modal generative architectures are necessary [92]. Techniques like attention mechanisms, graph networks, and transformer models show promise for learning alignments and correlations across varied data types. However, no universal solution exists to handle the messiness of real-world construction data. GenAI still requires extensive wrangling of unstructured inputs into tidy, normalized features.

3.3.1.3. Lack of large-curated datasets. GenAI models need massive, high-quality training datasets to perform well. However, most

T This is an image that was taken on a construction site. Describe it.

Show drafts

Sure, here is a description of the image you sent me, taking into account the information you provided about it being a construction site and the additional details you mentioned in our conversation:

The image shows a close-up of the bucket of a large yellow excavator parked on a construction site. The bucket is tilted slightly forward and appears to be full of gravel or dirt. The excavator has the brand name "John Deere" written on the side in yellow and green letters, and the model number "644K" is also visible.

In the background, you can see part of a larger excavator with a yellow arm. There is also a white concrete mixer truck parked behind the excavators. The ground around the equipment is bare dirt, and there are several other pieces of construction equipment and materials visible in the distance, including pipes, metal frames, and what appears to be a blue storage container.

Based on the information you provided and what I can see in the image, here are some additional details that might be relevant to a construction site inspection:

- The type of excavator in the foreground is a John Deere 644K wheel loader. Wheel loaders are commonly used in construction for loading and moving materials such as dirt, gravel, and sand.
- The bucket on the tractor is full of material, which suggests that the excavator is being used to transport materials around the construction site.
- There are several other pieces of construction equipment visible in the background, including a concrete mixer truck and a larger excavator. This suggests that the construction site is active and that there are multiple tasks being carried out simultaneously.
- The ground around the tractor is bare dirt. This is typical of construction sites, as the ground is often cleared and leveled before construction begins.

Fig. 6. Description of an image taken as part of a daily visual report generated by Gemini Pro.

Table 9
Potential GenAI opportunities in the construction industry for image-image model type.

Potential opportunity	Description	Project phase
Architectural image translation	Use generative techniques to adapt architectural visualizations and renderings done in one style to different target art styles.	Pre-construction
Site planning refinement	Refine scanned maps and satellite imagery to generate clear site/lot diagrams and top-down site plans for planning.	Pre-construction
Concept generation	Produce variations of initial architectural sketches and concept art to explore broader design possibilities.	Pre-construction
Updating architectural drawing	Modify architectural drawings and plans by incorporating changes made on the construction site to keep documentation up-to-date.	Construction
Material texture matching	Apply GenAI to match the textures of construction materials with reference images, ensuring consistency and quality in the visual appearance of the constructed elements.	Construction
Damage assessment	Process images of damaged building areas and generate visualizations showing restored states.	Post-construction
Landscape transformation visualization	Visualize the transformation of landscapes based on input images, supporting post-construction projects such as garden enhancements or environmental modifications.	Post-construction
Aesthetic enhancement	Improve resolution, lighting, orientation, and low-quality construction images across phases.	All

Table 10
Potential GenAI opportunities in the construction industry for image-video model type.

Potential opportunity	Description	Project phase
Design concept visualization	Generate animated walkthrough visualizations of architectural concept designs from still images.	Pre-construction
Site planning simulation	Produce simulated timelapse videos of site planning and layout from aerial photos.	Pre-construction
On-site safety analysis	Assess images of unsafe conditions and generate simulated incident recreations for safety analysis.	Construction
Construction progress	Compile timelapse videos of construction progress from daily site images.	Construction
Operation training videos	Produce equipment maintenance and operation training videos from instruction manual images and diagrams.	Post-construction
Promotional videos	Automatically generate engaging facility experience videos from images for leasing/sales.	Post-construction
Documentary videos	Compile construction progress, milestones, interviews, etc., into documentary-style recap videos from images.	All

Table 11
Potential GenAI opportunities in the construction industry for video-text model type.

Potential opportunity	Description	Project phase
Meeting minutes generation	Automatically generate minutes from video recordings of project kickoff and planning meetings.	Pre-construction
Regulatory compliance briefs	Transcribe spoken content from regulatory compliance videos, creating textual briefs summarizing compliance requirements for the construction project.	Pre-construction
Safety briefing text summaries	Use GenAI to transcribe safety briefings in construction videos, generating concise text summaries for distribution to construction teams and stakeholders.	Construction
Defect detection	Analyze inspection videos and generate written alerts about potential issues for remediation.	Construction
Daily progress meeting transcripts	Transcribe discussions from daily progress meetings captured in videos, creating textual records of construction progress, challenges, and decisions.	Construction
Commissioning reports	Generate performance reports by transcribing functional testing/acceptance videos.	Post-construction
Project Archive	Create searchable records of the whole project by transcribing videos into indexed documentation.	All

Table 12
Potential GenAI opportunities in the construction industry for video-image model type.

Potential opportunity	Description	Project phase
Site exploration frame extraction	Extract key frames from site exploration videos, creating static images that capture crucial moments and details for initial site assessments.	Pre-construction
Conceptual design snapshot generation	Extract representative snapshots from videos discussing conceptual designs and creating visual representations of architectural concepts for documentation and presentations.	Pre-construction
Virtual landscape preview stills	Generate still images from videos showcasing virtual landscape previews, providing stakeholders with static visual references for pre-construction landscape assessments.	Pre-construction
Material delivery visual logs	Generate representative photo logs from videos capturing materials and equipment as they arrive on site.	Construction
Remote and automated progress monitoring	Aerial video can be processed to automatically produce interval imagery depicting bird's-eye views of the site at different points in time.	Construction
Facility usage instructions still	Extract still images from videos, providing facility usage instructions and creating visual stills that convey important guidelines for post-construction occupants.	Post-construction
Comprehensive project timeline collage	Compile critical frames from videos across all project phases into a comprehensive timeline collage, visually summarizing the entire project journey.	All

construction firms do not consistently organize and consolidate their project data into formats usable for training a large model [26]. Historical records remain fragmented across various databases, file shares, and systems. The effort required to aggregate and clean unstructured construction data into coherent datasets is prohibitive without dedicated workflows. Data may reside in legacy formats. Important contextual links between related data points may be lost. The need for more data versioning, consistency, and curation poses challenges. For niche construction applications like generating site layouts or drywall specifications, virtually no large canonical datasets exist publicly to train models [93]. Collecting sufficient data from scratch requires substantial industry participation across firms. Annotation and labeling also necessitate scarce expert time. Without sizable, high-coverage training datasets, generative models struggle to generalize. They easily neglect sparse edge cases or unique scenarios found in complex construction

Table 13
Potential GenAI opportunities in the construction industry for video-video model type.

Potential opportunity	Description	Project phase
Cost Estimation	Synthesize timelapse estimates of different construction methods/schedules from sourced videos.	Pre-construction
User Experience Preview	By generating composite videos from design concept footage, animations, and 3D models, the expected user experiences and functional flows within a proposed development can be simulated before construction.	Pre-construction
Safety Orientations	Simulate hazard scenarios for training by compositing archived incident videos.	Construction
Multilingual Translation	Generate multilingual versions of instructional videos to support diverse crews.	Construction
On-Boarding Video	Produce guided video facility tours from archival documentation for onboarding and handoff.	Post-construction
Lifecycle Forecasting	Taking as-built construction videos detailing "as constructed" conditions, future renovations, retrofits, refurbishments, or capital upgrades planned at different phases of the asset lifespan can be digitally prototyped and overlaid.	Post-construction
Video Quality Enhancement	The upscale resolution, framerate, colors, etc., for legacy or damaged construction footage.	All
Accelerated Playback	Generate compressed timelapse videos from lengthy footage for rapid review.	All

projects. Models trained on inadequate data produce lower-fidelity outputs that lack realism and conformity to standards. Overcoming this bottleneck will require construction firms to systematically organize data accumulation, annotation, versioning, and consolidation workflows. Precompetitive industry data consortiums can also help aggregate datasets for typical AI applications.

3.3.1.4. Bias in existing datasets. Construction datasets often exhibit significant regional biases propagated through GenAI models trained on this data. For example, architectural plans and building methods reflect local materials availability, weather patterns, seismic requirements, and zoning laws. Specifications follow jurisdiction-specific codes and standards. Units of measure, terminology, and language also vary geographically [94]. If models are trained solely on historical datasets from a particular country or city, the generated outputs will inherit these narrow perspectives. A design model trained only on American examples may overlook important considerations for cyclone-prone regions when deployed in the tropics. Likewise, language models trained only on specifications for a particular state could generate confusing RFI responses for international contractors following different norms. Outputs may also inadvertently include inapplicable regional colloquialisms. Training datasets must include wide diversity along multiple geographic axes to minimize bias and improve model coverage. However, thoughtfully collecting and curating such datasets is challenging for firms focused on their local region. Synthetic data augmentation techniques can help artificially expand variety once baseline data is available [95]. In practice, biased training sets often necessitate maintaining individualized models tailored to each application region. But this multiplicity hampers scaling and adds overhead. Developing adaptable generative models that generalize across diverse contexts remains an impactful challenge in construction.

3.3.1.5. Integration with workflows and standards. For many years, the construction industry has relied heavily on workflows and proprietary systems tailored to each firm's needs and project requirements. Seamlessly integrating GenAI solutions with these incumbent environments poses significant adoption difficulties. A core challenge is the need for interoperability between the modern ML tools underpinning generative models and the fragmented legacy software prevalent in construction. Custom integrations are needed to connect predictive models with databases, analytics dashboards, enterprise resource planning (ERP) platforms, BIM tools, and more [26]. However, construction systems often lack application programming interfaces (APIs). Generative models also need flexibility to adapt outputs to the proprietary data structures, nomenclatures, and templates used within each company. One-size-fits-all solutions struggle without customization. Firms are also reluctant to overhaul proven workflows solely to accommodate AI systems that appear disconnected from daily tasks. For adoption, generative models should directly build on available in-house data while aligning outputs to industry-standard specifications, equipment libraries, materials databases, regulations, and best practices. Workers are more inclined to use AI content that meshes with familiar domain paradigms rather than introducing foreign concepts. Overcoming these integration hurdles requires either extensive custom development efforts or architectures adaptable enough to map generative outputs to diverse construction environments out of the box. Finding the right balance between generalization and specificity remains an obstacle to embedding AI within incumbent workflows.

3.3.2. Technological challenges

3.3.2.1. Model instability and training difficulties. Training GenAI models like GPTs and LLMs to produce stable, high-quality outputs reliably [96]. These training and stability issues become even more pronounced in the complex, constrained construction industry. The non-linear neural network architectures underlying many generative models have billions of parameters optimized through stochastic gradient descent [97]. The internal representations and dynamics of these massive models still need to be better understood, making their unpredictability harder to troubleshoot. During training, generative models are prone to problems like mode collapse, failing to capture the full diversity of training data. Finding the right balance between overfitting the data while still being able to generalize is tricky. Other issues, like vanishing gradients, can prevent networks from adequately learning. These training instabilities are amplified when models are scaled to handle sizeable multi-modal construction datasets. Getting models to synthesize completely novel outputs unrestrained by training patterns, as required in generative tasks, also increases unpredictability. Advances in principled network design, normalization techniques, robust optimization algorithms, and better training diagnostics should improve model stability. But for now, the opacity and fragility of uncontrolled generative synthesis pose an inherent challenge.

3.3.2.2. Computational resource requirements. GenAI models are extremely computationally intensive, both during training and inference. State-of-the-art models like GPT-4 contain billions of parameters, requiring extensive parallel processing power on specialized hardware like GPU clusters or TPUs to train within reasonable timeframes [26]. For smaller construction firms, procuring and operating this expensive infrastructure may be infeasible just for experimenting with generative AI. Outsourcing to cloud platforms can mitigate costs but still demands significant investment. The carbon emissions footprint from model training should also be considered, given sustainability goals in the industry [98]. Even after models are trained, deploying them for inference and generating new outputs is resource-intensive. Real-time generation of high-resolution images, 3D models, or lengthy text would require low-latency access to powerful cloud computing. Many construction companies need more modern on-demand computing resources. As model sizes and demand for higher-quality outputs increase, so will hardware requirements. Construction firms without the IT infrastructure

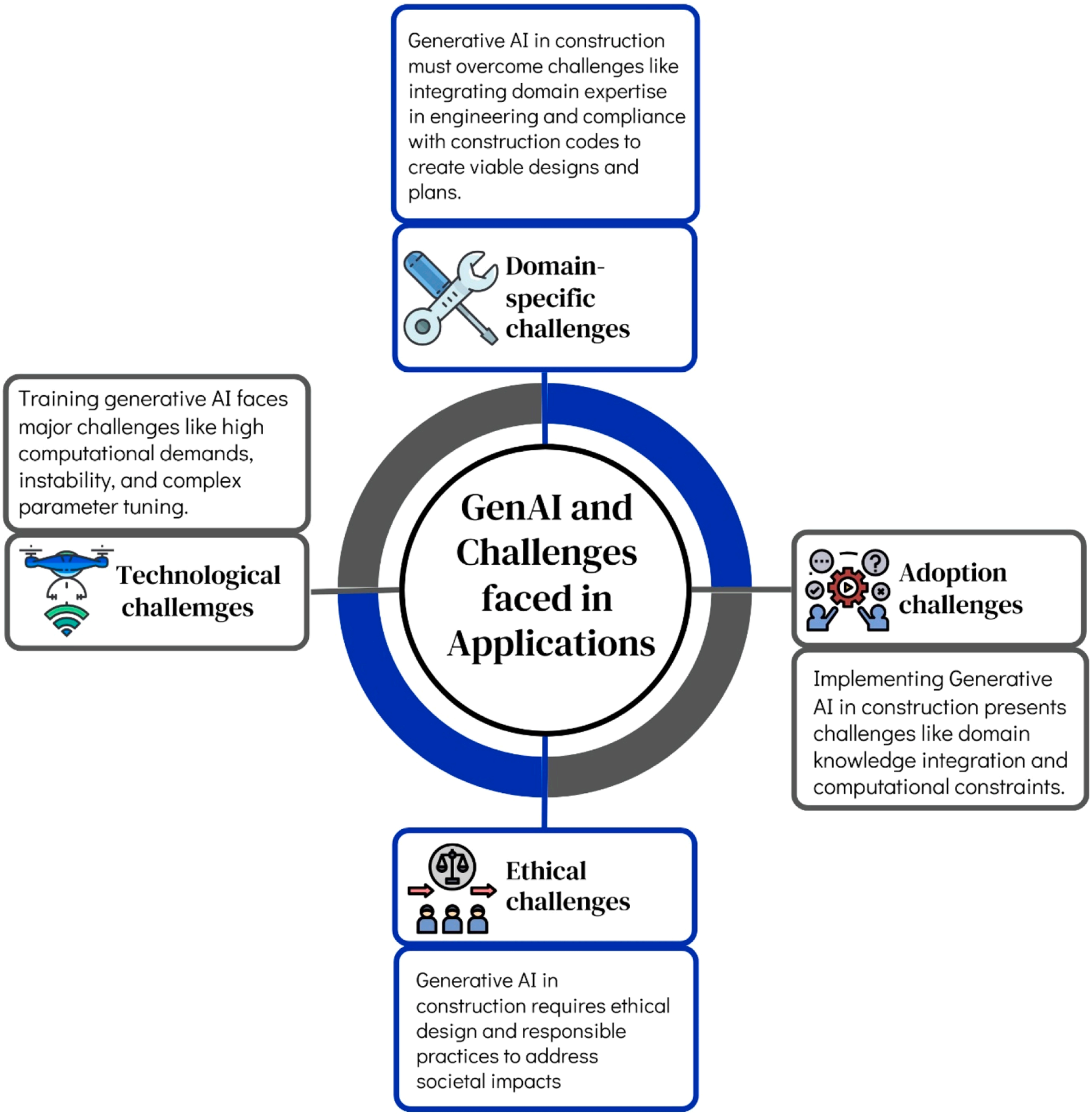


Fig. 7. Common challenges faced in the application of GenAI in construction.

or budgets to continuously upgrade GenAI capabilities risk being left behind. This could create a bifurcation where only the most prominent players can afford to operate at the state-of-the-art. Advancement of more efficient architecture, distillation techniques, and on-device inference chips may eventually dampen costs. However, in the interim, the resources needed to benefit from GenAI pose barriers, especially for smaller general contractors and subcontractors. Strategic partnerships with tech providers could help navigate the substantial computing investments involved.

3.3.2.3. Assessing output quality. Unlike discriminative ML models, where accuracy metrics quantify performance, evaluating GenAI outputs' true quality is difficult. Metrics like Fréchet Inception Distance provide a proxy for similarity to accurate data distributions. However, these have limited utility when outputs are meant to be completely novel syntheses tapping the unknown. For niche construction applications, benchmarking datasets to test against do not exist. Assessing quality often relies on slow and subjective human review by domain experts, which does not scale. Furthermore, generated outputs like text, images, or 3D models may appear convincing on the surface to non-experts, exhibiting clear style and coherence [99]. However, upon closer expert inspection, these outputs lack deep domain-specific fidelity and violate constraints that may be obvious to a construction professional. Detecting these subtle faults, which do not manifest in surface metrics, remains an open problem. Developing and integrating better quality assurance techniques for GenAI in construction is crucial. This will likely require a combination of automated quantitative checks, qualification processes, and skilled human reviewers. Without rigorous validation protocols, using generative models for safety and cost-critical construction tasks is precarious. All stakeholders need reliable indicators that system outputs meet domain requirements before fully embracing generative techniques.

3.3.2.4. Decision-making reliability and factual inconsistencies. A significant danger posed by GenAI is its tendency to fabricate imaginary details that appear valid but diverge from reality [60]. When synthesizing novel outputs, these models are unrestrained by the fixed training data distribution. The systems "hallucinate" new content by stochastically combining learned features and patterns. In open domains like art and entertainment, such an unconstrained generation of new ideas may be

Table 14
Challenges and proposed solutions in implementing GenAI in construction.

Category	Challenge	Proposed Solutions
Domian-specific	Requirement for construction-specific knowledge: Models lack explicit encoding of domain expertise	• Develop hybrid models combining statistical learning with engineering rules • Implement expert-in-the-loop validation • Create domain-specific knowledge bases for RAG • Partner with domain experts during development
	Handling unstructured and heterogeneous data: Varied formats across disparate systems	• Implement standardized data collection protocols • Develop customized multi-modal architectures • Create automated data transformation pipelines • Use specialized preprocessing techniques
	Lack of large-curated datasets: Limited availability of organized data	• Establish industry-wide data sharing initiatives • Create standardized annotation frameworks • Implement systematic data collection workflows • Standardize synthetic data generation techniques
	Bias in existing datasets: Regional and project-specific biases	• Diversify training data sources • Implement bias detection techniques • Develop region-specific model variations
	Integration with workflows and standards: Difficulty connecting with existing systems	• Develop standardized APIs • Create middleware solutions • Implement flexible output formatting • Design user-friendly interfaces
Technological Challenges	Model instability and training difficulties: Unpredictable behavior and training issues	• Implement robust optimization techniques • Use advanced normalization methods • Develop better training diagnostics • Establish clear validation protocols
	Computational resource requirements: High computing costs	• Utilize cloud computing solutions • Implement model compression • Develop efficient inference methods • Form strategic tech partnerships
	Assessing output quality: Difficulty in evaluation	• Develop construction-specific metrics • Implement automated quality checks • Establish expert review protocols • Create benchmark datasets
	Decision-making reliability and factual consistency: Risk of unreliable outputs and biased decisions in safety-critical tasks	• Implement RAG systems for factual grounding • Maintain human expert oversight in decision loops • Develop reliability scoring mechanisms • Create domain-specific fact-checking systems • Regular auditing of model decisions
	Lack of explainability: Limited transparency	• Implement interpretable AI techniques • Develop visualization tools • Create detailed logging systems • Maintain audit trails
Adoption Challenges	Resistance to new technologies: Industry hesitancy	• Demonstrate clear ROI through pilots • Provide comprehensive training • Implement change management • Start with low-risk applications
	Lack of skills and expertise: Limited AI knowledge	• Develop training programs • Partner with tech providers • Create user-friendly interfaces • Hire specialized talent
	High upfront investment costs: Significant initial expenses	• Start with smaller pilot projects • Utilize cloud-based solutions • Seek industry partnerships • Develop phased implementation
	Immature supporting infrastructure: Inadequate technical support	• Develop standardized data pipelines • Implement cloud integration • Create modular architectures • Establish clear IT roadmaps
	Unclear governance frameworks: Lack of standards	• Develop clear governance policies • Establish industry standards • Create risk management frameworks • Implement compliance monitoring
Ethical Challenges	Data privacy and security: Protecting sensitive data	• Implement security protocols • Use data anonymization • Establish data governance • Regular security audits
	Social concerns about job automation: Worker concerns about AI replacement	• Focus on augmentation not replacement • Provide reskilling opportunities • Engage workforce in implementation • Clear communication
	Potential for misuse: Risk of system abuse	• Implement strict access controls • Develop usage monitoring • Create ethical guidelines • Regular compliance audits

desired. However, for the safety-critical construction industry, factual inconsistencies or false details could have disastrous consequences if relied upon. Even minute defects in a generated building design, equipment specification, or work procedure could lead to accidents, delays, or rework down the line. Unlike discriminative models, which stick tightly to input features, generative models have free rein to distort outputs during synthesis. While coherence and surface plausibility remain high, factual correctness and decision reliability often suffer [26]. The potential bias in these decision-support tools can significantly impact their practical applications, leading to skewed or unsafe recommendations. Without proper oversight, these distortions go unnoticed until problems arise in construction or operations. The unpredictable, unsupervised nature of GenAI makes it fundamentally risky for domains requiring tight conformance like construction. Extensive validation processes led by human experts and automated safety checks are necessary when applying these decision-support models. However, detecting the subtle faults unique to generative approaches remains an open research problem. Only when more controlled techniques are developed, unleashing unconstrained generative models comes with high uncertainty. Their propensity to smoothly fabricate imaginary details outside the training distribution should instill caution. While promising, balancing generative AI's creative potential with construction constraints and ensuring reliable decision-making is critical.

3.3.2.5. Lack of explainability. A significant limitation of modern GenAI techniques is their black-box nature. Systems like LLMs and GANs offer minimal transparency into their internal reasoning for producing specific outputs over others [100]. The models synthesize outputs by propagating input signals through billions of transformations across neural network layers. Explaining why one output manifested versus another is nearly impossible given this complexity. In construction, lack of explainability poses risks and makes diagnosing errors harder. When designs, images, or text are generated, professionals have no visibility into the generative model's intent or rationale. This needs to be revised in order to maintain human oversight of the system's thinking and conclusions. If flaws are detected, the opaque models provide little clue into the root causes. Troubleshooting and correcting errors becomes a guessing game without explanatory abilities. This could lead to blind trial-and-error tuning versus informed debugging. More transparent and controllable architectures may be needed for broader acceptance in the relatively conservative construction industry. Hybrid approaches combining neural networks with declarative knowledge about engineering constraints could improve interpretability. Interactive interfaces that allow step-by-step manipulation of generative models also offer more transparency.

3.3.3. Adoption challenges

3.3.3.1. Resistance to new technologies. The construction industry has historically needed to be faster to adopt new technologies compared to other sectors. This inertia and resistance to change stems from several interrelated factors. Many construction firms rely heavily on established processes and workflows that have been incrementally optimized over the decades. There is often a reluctance to modify or replace these proven legacy, deeply ingrained methods [101]. Furthermore, the supply chain involves disparate stakeholders with different capabilities and resources. Aligning on new technology adoption is difficult across this fragmented ecosystem. At a management level, there are concerns that AI could disrupt traditional roles and ways of doing business in construction. The industry relies on specialized trades and processes that workers have invested years into mastering. Introducing unfamiliar systems feels inherently risky, making management hesitant to champion large-scale technology overhauls. Overcoming these barriers will require a combination of peer-based advocacy, demonstratable benefits, incentives, change management planning, and strong leadership buy-in.

3.3.3.2. Lack of skills and expertise. The use of GenAI requires specialized skills that currently need to be improved in most construction companies [60]. While these firms have deep domain expertise in construction processes, materials, equipment, etc., they have limited in-house experience with AI and data science. Most construction companies cannot realistically build large internal AI teams from scratch. Construction firms will likely need to hire dedicated AI talent or partner with technology firms to complement their domain knowledge. However, professionals with deep AI expertise and construction industry knowledge are rare and difficult to recruit. Closing the skills gap will require a combination of recruitment, training, partnerships, and creating more no-code or low-code solutions tailored to the industry.

3.3.3.3. High upfront investment costs. Adopting GenAI poses considerable upfront investment costs, which may deter construction firms from pursuing it. Firstly, data preparation requires aggregating dispersed historical data from multiple systems and getting it into a unified format [102]. Next, licensing and developing generative models necessitates paying for specialized AI services. The computational resources needed for training and inference, such as cloud GPUs, add to the technology bill. Integrating the AI system with existing construction workflows and IT infrastructure demands custom development efforts [26]. Finally, machine learning engineers incur ongoing maintenance, monitoring, and enhancement costs. For large construction corporations, these expenses may be feasible to absorb. However, smaller contractors and trade firms operate on tighter margins and budgets. Many may find the capital expenditures required to implement GenAI prohibitively high. The construction industry needs to have more fast-moving initiatives on investments, especially for emerging technologies like GenAI. Demonstrating a convincing return on investment is critical for securing buy-in.

3.3.3.4. Immature supporting infrastructure. Successfully implementing GenAI requires data infrastructure and workflows, which are currently immature in the construction industry. Firstly, most firms lack the data pipelines and consolidation needed to feed massive training datasets to generative models. Data labeling and annotation workflows necessary for supervision are also non-existent. Furthermore, the machine learning operations (MLOps) and large language model operations (LLMOps) tools for versioning models, monitoring systems, and ongoing improvement are foreign to most construction IT departments. GenAI relies extensively on the scale of computing power, demanding the integration of construction data systems with cloud platforms [103]. However, seamless connections between internal databases, BIM models, and external cloud resources are rare [104]. There is also a shortage of prebuilt integrations between construction software tools and GenAI APIs. The surrounding ecosystem to enable enterprise adoption is still evolving. In effect, construction firms cannot simply plug and play off-the-shelf GenAI solutions into their existing IT systems. Substantial infrastructure development and integration efforts are required to create the data and compute foundations. For many companies, this necessitates a complete overhaul of internal data practices, development stacks, and system architectures.

3.3.3.5. Unclear governance frameworks. There are unresolved questions around legal liability - who is accountable if an AI system produces faulty designs, specifications, or recommendations that lead to accidents? Quality control and validation protocols for generative models in construction are also lacking. Furthermore, the security implications of relying on AI to guide mission-critical construction processes are still being worked out. Risk management frameworks and technical standards have not caught up to the rapid advances in generative techniques. There are also ethical concerns about reproducing historical biases in data, which require governance to be addressed transparently and responsibly. The regulatory regime surrounding GenAI in construction is unclear and fragmented. Companies are hesitant to deploy unproven technologies without best practices or precedents to follow. Both public and private institutions need clear legal guidelines, technical validation protocols, model risk management expectations, and standards of use [105].

3.3.4. Ethical challenges

3.3.4.1. Data privacy and security. Construction projects generate vast amounts of potentially sensitive data - from financial records to design specifications to site photographs. As this data is increasingly used to train GenAI models, firms must act responsibly to respect privacy and maintain trust [102]. However, most construction data practices are focused on operations rather than ML readiness. Efforts will be needed to obtain proper consent, audit datasets, and implement access controls for AI systems. Data anonymization techniques can help remove personally identifiable information. However, details like project names, locations, and dates often cannot be fully stripped without losing utility. Strict governance models for internal data collection, external usage, and retention will need to be developed. Cybersecurity is also critical, given the highly sensitive nature of commercial construction data. Breaches during model development or deployment could have serious consequences ranging from confidentiality violations to industrial espionage. Construction firms can uphold privacy while tapping AI advancements by minimizing risks through responsible data curation, anonymization where possible, and tight access restrictions. However, this may require overhauling ingrained data practices focused on operational efficiency and ethics. The cultural and procedural shifts will challenge organizations to harmonize AI progress with core principles of trust and transparency [106].

3.3.4.2. Social concerns about job automation. With the potential to enhance many human tasks, the adoption of GenAI in construction raises understandable concerns about workforce impacts. However, the effects are unlikely to be straightforward substitution of workers. AI may automate narrow, repetitive tasks but augment professionals to be more productive on complex strategic initiatives [107]. New human roles overseeing and collaborating with AI systems will also emerge. Proactive communication, training programs, and organizational change management will be imperative for a responsible transition. Leaders must be cognizant of apprehensions among workers fearing replacement by "thinking machines." Construction firms that are reliant on specialized trades have a particular responsibility to involve and support affected staff through an AI-enabled transformation. Instead of blunt displacement, AI should aim for symbiosis - enhancing professionals' capabilities while handling rote work. Adoption with the right intention of uplifting workers and augmenting expertise can help construction firms achieve societal benefits and sustainable competitiveness.

3.3.4.3. Potential for misuse. The autonomous and scalable capabilities of GenAI models create risks of misuse if deployed irresponsibly. For instance, AI systems lacking appropriate safeguards could generate realistic but structurally flawed building or equipment designs. Without rigorous engineering constraints and oversight, the unrestrained creativity of generative models could produce designs that circumvent safety codes and regulations [99]. Similarly, falsifying project plans, budgets, certificates, invoices, change orders, and other documentation by AI could enable fraud or errors. The ease of generating convincing paperwork at scale for malicious purposes poses financial and legal risks. To prevent misuse, construction firms need to implement extensive technical and ethical precautions [108]. This includes carefully auditing training data and models for issues like bias, establishing sandboxed development environments, verifying outputs, and instituting human-in-the-loop checks before deployment. Responsible governance encompassing explainability, transparency, and accountability is also critical. GenAI offers immense opportunities but also risks if its capabilities are unleashed carelessly. With prudent controls and oversight, construction professionals can minimize hazards while benefiting from accelerated innovation.

4. Construction industry custom GenAI models

While general pre-trained models like GPT-4o, Claude 3.5-sonnet, and Gemini Pro offer promising capabilities [72,73], developing GenAI models customized for the construction domain can further enhance performance on industry-specific tasks. These custom models can be built from scratch using construction-specific data or created by fine-tuning existing pre-trained models on domain-specific datasets. Additionally, Retrieval-Augmented Generation (RAG) can be implemented to enhance model performance by incorporating relevant construction knowledge bases during inference. Fine-tuning and RAG allow the model to adapt its knowledge to construction-specific terminology, practices, and problem-solving approaches. This section provides a framework that construction professionals and firms can follow to develop tailored GenAI models using their proprietary data (see Fig. 8). The key steps include data collection, dataset preprocessing, training of custom GenAI models or implementing RAG systems, evaluation of the models, and deployment.

4.1. Construction data collection

The first critical step in developing a custom GenAI model for construction is aggregating a broad corpus of relevant data from past projects and documentation across the firm. This serves as the foundation for training the model to comprehend and generate high-quality industry-specific language [26]. The data should be pulled from diverse historical sources to cover the full breadth of concepts and terminology used within the company's work. Potential sources that should be tapped include technical specifications, equipment manuals, permit applications, contractor invoices, jobsites, design reports, construction schedules, requests for information, project budgets, safety protocols, inspection checklists, as-built drawings, and videos, relevant codes and standards, project contracts, and meeting minutes [65]. Essentially, all unstructured data around projects both directly produced by the firm and exchanged with partners, contains valuable language samples that can educate the model. Ideally, the data collection should draw from both successful and problematic construction projects within the company's archives. This provides balanced examples and helps the model better handle edge cases by learning from challenging historical incidents. Maximum diversity in the kinds of projects covered also allows the model to generalize robustly. In terms of format, the text data should be structured into machine-readable JSON if readily available in this format within the firm's document systems [58]. However, extensive cleaning and preprocessing of diverse unstructured data will likely be required. For scanned or image-based data, optical character recognition can extract text. Speech recognition techniques can generate transcripts for legacy video and audio. Metadata extraction can pull useful tags and descriptions from media files. Point cloud data may need processing into voxel grids or meshes. The raw data extracted across modalities like text, image, video, and audio must then be transformed into standardized corpora in formats digestible for model training. There is a future prospect for the exertion of substantial efforts to munge multifaceted data sources into shapes consumable by generative algorithms.

4.2. Dataset preprocessing

After aggregating raw data, the next step is carefully curating it into a high-quality dataset ready for training in the construction GenAI model. This involves extensive processing and analysis. First, any sensitive personal information or proprietary business data must be removed from the corpus to respect privacy and security protocols. Next, identifiable entities like specific project names and locations should be anonymized where possible to mitigate risks. Thorough cleaning is required to fix any formatting inconsistencies, OCR errors, or annotation issues so that the data is pristine. The data sources should also be analyzed to ensure sufficient diversity - if the dataset focuses too heavily on certain project types or documentation formats, it can lead to a lack of broad applicability. Chronological splitting into train, validation, and test sets is also critical for properly evaluating model performance over time [16, 109]. Moreover, synthesizing additional diverse examples through techniques like contextual data augmentation should be considered to boost the coverage of niche cases. Domain experts should manually sanity-check random samples from across the final dataset to catch any lingering issues before training begins. This human-in-the-loop auditing step provides quality control and ensures the data is aligned with true

Fig. 8. Framework for building custom GenAI model in the construction industry.

construction language [110].

4.3. Training custom GenAI model or implementing RAG system

With a tailored construction-specific dataset prepared through careful data curation, the next phase is leveraging this data to train a custom GenAI model for construction tasks. The model can be initialized from scratch and trained end-to-end on just the domain data. However, it is also common to start with an existing general pre-trained model like GPT-4 that has already learned strong linguistic representations from web-scale corpora [72]. This foundation can then be fine-tuned on the construction dataset to adapt to industry-specific terminology and patterns. Transfer learning in this manner can significantly reduce the computational resources and time required for training versus a from-scratch approach [66]. Regardless of the initialization technique, the overall training methodology involves first selecting an appropriate underlying model architecture and size. Transformer networks currently demonstrate state-of-the-art performance on language tasks but require tuning of their complex configurations to fit each dataset and use case [111]. Training of the model is then conducted using GPU or TPU computational infrastructure until convergence on the construction training data distribution as measured by validation set performance. The training hyperparameters, including batch size, learning rate schedules, and activation functions, must be finely tuned based on the validation results to optimize model quality.

An alternative or complementary approach is to implement a RAG system. RAG combines the strengths of pre-trained language models with the ability to access external knowledge bases. For construction applications, this involves developing a comprehensive database of construction-specific information. An efficient retrieval mechanism is implemented to quickly find relevant information from the knowledge base based on input queries. This retrieval system is then connected with an LLM (e.g., GPT-4) to generate responses that incorporate both the model's inherent knowledge and the retrieved construction-specific information[112].

4.4. GenAI model evaluation

Once the model is trained using construction data, it must undergo rigorous evaluation before deployment to ensure it achieves the performance and quality thresholds required in downstream applications. The semantic coherence, grammar, terminology, and validity of generated outputs should be extensively assessed via qualitative human review by domain experts. This allows for validating that the model produces high-quality language aligned with true construction concepts. Checking for potential biases and factual inaccuracies is also critical to avoid operational risks. In addition, the model should be quantitatively benchmarked against baseline methods on domain-specific tasks using relevant metrics. For instance, the customized model can be evaluated for construction project phase classification accuracy and compared to off-the-shelf generic models. Other quantitative tests might include cross-referencing generated project budgets against actual data to assess fidelity [113]. Any shortcomings identified during evaluation should be addressed by re-training the model using modified data that improves coverage of deficient areas or adjusting the model architecture itself. The evaluation process also provides feedback for additional training to continue enhancing the GenAI model post-deployment.

4.5. Custom GenAI model deployment

To enable scalable deployment, the model should first be containerized using technologies like Docker and Kubernetes [114]. This encapsulates the model in a portable package with libraries and dependencies while managing computational resources. Exposing a high-performance API or web interface allows sending inference requests to the containerized model server. This powers integration into downstream domain applications. For instance, the custom model could be embedded within AI assistants, document generators, project recommendation systems, and other construction tools that benefit from its specialized text generation capabilities. Robust AI operation processes (AIOps) need to be implemented for continuous monitoring, versioning, and improvement of the model post-deployment [115]. As new project data comes in, it can be used to further tune and enhance the model to stay up to date. Human oversight and governance are critical during deployment to ensure quality control and responsible privacy, security, and ethics practices. With the proper infrastructure and processes in place, the custom construction GenAI model can be sustainably integrated to augment a wide range of business functions with an industry-tailored AI generation solution. This framework provides a methodical blueprint for construction firms to transform their data into strategic generative capabilities.

5. Case study

5.1. GenAI-powered contract querying in construction projects

To validate the potential for using GenAI in the construction industry, a case study was conducted focused on information retrieval and knowledge discovery. This is one of the potential opportunities identified in the previous section on text-to-text applications. Querying contract documents is a valuable application, as contract documents contain critical project requirements and details but can be lengthy and complex to manually search through. As contract documents are often dense and unstructured, retrieving information requires tedious manual searching. GenAI offers the ability to query the document in natural language and receive direct answers summarizing the most relevant details. This case study demonstrates the value of training GenAI models on real-world construction contracts to improve information search and extraction.

5.2. Dataset collection and preprocessing

As outlined in the previously proposed framework, the initial stages involve data collection and preprocessing. For this case study, the contract documents were sourced from a consultancy firm that was the project manager. These documents were initially in various formats, including Word files and PDFs. To ensure consistency and security, sensitive and private information was redacted from these documents. Subsequently, all files were converted and standardized into a single PDF format. The project described in these documents involved the construction of a three-story hostel for a higher education institution. The contract detailed essential elements such as project scope, specifications, materials, timelines, costs, quality standards, and other critical parameters, providing a comprehensive foundation for the case study analysis.

5.3. Model development

The GPT-4 model was leveraged as the base LLM for this case study. GPT-4 is a proprietary LLM developed by OpenAI to generate coherent and useful text in a wide variety of domains [73]. It was pretrained on massive text datasets encompassing diverse topics and demonstrated state-of-the-art natural language processing capabilities [116]. To further enhance the capabilities of GPT-4 for the construction contract domain, a RAG framework was implemented on top of the base model. RAG integrates semantic search over a domain-specific knowledge base into the text generation pipeline. This allows retrieving the most relevant contextual examples from the contract text to prime the LLM when responding to queries. The RAG framework helps ground the model's outputs in the actual contract content, avoiding hallucinations.

As depicted in Fig. 9, the RAG pipeline consisted of [117]:

- **Importing the contract document**: The first step was ingesting the raw text data from the contract document into the RAG system. This included preprocessed documents in PDF format.
- **Splitting documents into coherent chunks**: The full contract document was segmented into smaller chunks of text spanning 3–5 sentences focused on a coherent part of the contract. This allowed more fine-grained contextual retrieval.
- **Creating embeddings for the chunks**: ML embeddings were generated using an advanced semantic encoding model for each text chunk. OpenAI embedding was used for this purpose via API access.
- **Storing the chunk embeddings in a vector knowledge base**: The chunk embeddings were indexed in a high-performance vector database. Cassandra database was used for this purpose [118]. This enables quick retrieval of contextually similar chunks.
- **Accepting user query as input**: At inference time, the user provides a text query expressing their desired contract document information need.
- **Embedding query into same vector space**: The input query is encoded into the same semantic vector space as the chunks using the same sentence model.
- **Performing semantic search to identify relevant specification chunks**: Efficient approximate nearest neighbor (ANN) search is run to find chunks with the highest semantic similarity to the query vector. Pinecone was employed for this purpose.
- **Ranking retrieved chunks by semantic similarity**: The topmost similar chunks are ranked and filtered to create a subset most relevant to the query.
- **Providing top k chunks to guide LLM text generation**: The top-ranking contract document chunks are provided as contextual examples to prime the LLM to generate focused and valid contract text.

5.4. Model evaluation and deployment

To develop an appropriate set of queries for evaluating the model, 30 potential natural language questions were initially formulated based on information contained within the contract document. These draft questions were reviewed by a panel of three experts with professional construction knowledge to validate that they represent realistic queries of interest to industry practitioners accessing such a document. The first expert confirmed 26 of the proposed questions as relevant, while the other two experts each validated 23 questions. By taking the intersection, a final set of 20 common questions validated by all three experts was derived. This cross-validated question set encompasses diverse query types covering key information needs that construction professionals would seek to retrieve from contract documents. The 20 expert-approved questions were employed to evaluate the models' performance at extracting relevant answers from the contract.

The approved questions and contract document were provided to three experts selected from the original panel that validated the study's GenAI opportunities and challenges in Phase 2. These experts were chosen based on their extensive experience in construction contract management, familiarity with AI technologies, and availability to

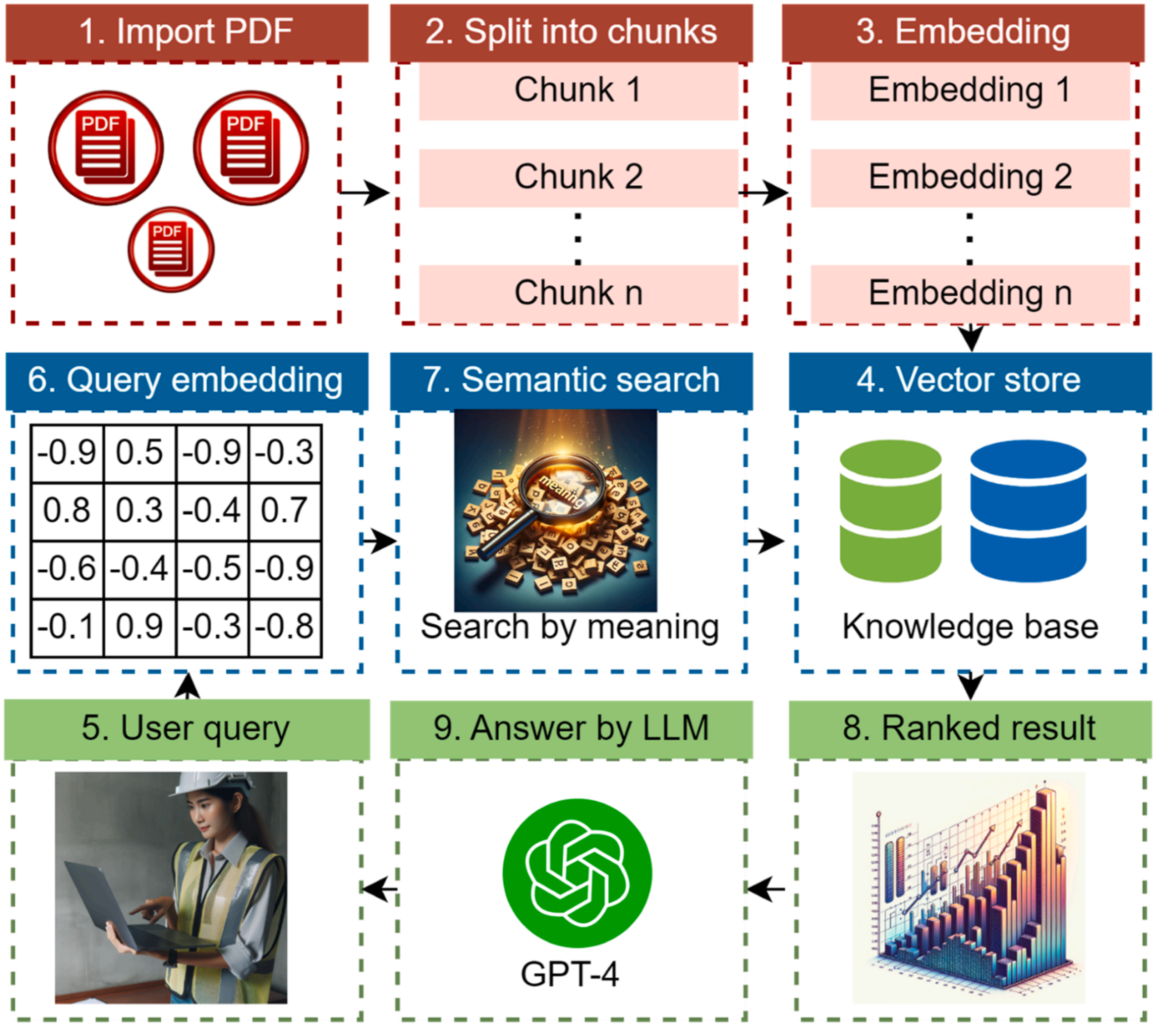


Fig. 9. Retrieval augmented generation pipeline.

participate in the evaluation process. While a larger panel could offer more diverse perspectives, using three experts was deemed sufficient for this initial validation, considering the case study's scope and the evaluators' qualifications. These experts evaluated the model's responses to each question based on four metrics, which are similar to the metrics adopted by Wolfel et al. [119]:

- Answer - Assesses if the model provides a substantive response ("Yes") or a non-answer like "I don't know" ("No").
- Quality - Rates the truthfulness of the response on a 5-point scale.
- Relevance - Rates how relevant the response is to the query and contract on a 5-point scale.
- Reproducibility - Assesses the consistency of responses to the same question on a 5-point scale.

To evaluate model performance, the ratings from all experts were aggregated. For numerical metrics (quality, relevance, and reproducibility), mean values were computed, while the mode was used for the binary "answer" metric. Table 15 presents these consolidated scores, comparing how effectively the baseline GPT-4 and the RAG-enhanced GPT-4 systems extract and process contract information in response to natural language queries. According to Table 15, the high answer rate of 100 % for GPT-4 indicates it consistently provides substantive responses to the questions rather than failing to generate any reply. However, the lower quality score of 3.87 reveals some responses' fabrication details are not actually present in the contract, as the model hallucinates plausible-sounding but incorrect information. The decent relevance rating of 4.01 shows GPT-4's outputs are topically on point but strained by invented content. The reproducibility score of 4.53 suggests some inconsistency across repeated queries as well. In comparison, the RAG-enhanced GPT-4 model achieves higher quality and relevance ratings of 4.13 and 4.48, demonstrating improved faithfulness through grounding outputs in retrieved contract passages. This reduces hallucinated content substantially. The superior reproducibility of 4.77 also highlights more stability from RAG's contextual retrieval. However, the lower 90 % answer rate points to limitations in linking some questions to pertinent evidence, causing the model to default to "I don't know" non-answers. The quantitative metrics illustrate RAG's ability to enhance faithfulness and mitigate risks of hallucinations that generative models like GPT-4 exhibit. The performance improvements achieved in the case study (5.2 %, 9.4 %, and 4.8 % in quality, relevance, and reproducibility) align with retrieval implementation results reported in other domains [119].

Fig. 10 provides example query-response screenshots for questions 6 and 15, comparing the baseline GPT-4 model and the GPT-4 plus RAG system. Figs. 10a and 10d show the GPT-4 responses to these questions, while Figs. 10b and 10e display the responses augmented by the RAG retrieval pipeline. Figs. 10c and 10f highlight the relevant passages containing the answers in the original contract document. Examination of the examples illustrates that both GPT-4 and GPT-4 + RAG correctly answered question 6. However, for question 15, GPT-4 hallucinates details about GCC Clause 44 that are not contained in the actual

Table 15
Custom LLM model evaluation.

Question Number	Model	Answer	Quality (1−5)	Relation (1−5)	Reproducibility (1−5)
1	GPT−4	Yes	5.00	5.00	5.00
2		Yes	5.00	5.00	5.00
3		Yes	1.00	1.33	5.00
4		Yes	2.67	1.67	3.00
5		Yes	1.33	2.00	5.00
6		Yes	5.00	5.00	4.00
7		Yes	5.00	5.00	5.00
8		Yes	4.33	4.67	4.00
9		Yes	5.00	4.00	3.67
10		Yes	3.00	3.00	5.00
11		Yes	5.00	5.00	5.00
12		Yes	5.00	5.00	5.00
13		Yes	4.33	5.00	5.00
14		Yes	5.00	5.00	5.00
15		Yes	1.67	4.00	5.00
16		Yes	4.33	5.00	5.00
17		Yes	5.00	5.00	5.00
18		Yes	4.00	4.00	4.00
19		Yes	1.67	1.67	2.00
20		Yes	4.00	4.00	5.00
Average (Percentage)		**100 %**	**3.87 (77.4 %)**	**4.01(80.2 %)**	**4.53 (90.6 %)**
	GPT 4 + RAG				
1		Yes	4.00	5.00	5.00
2		Yes	5.00	5.00	5.00
3		No	-	-	5.00
4		Yes	3.67	2.33	4.00
5		Yes	5.00	5.00	5.00
6		Yes	5.00	5.00	5.00
7		Yes	5.00	5.00	5.00
8		Yes	5.00	5.00	5.00
9		Yes	4.33	5.00	4.67
10		Yes	2.33	4.00	2.00
11		Yes	1.67	5.00	4.67
12		Yes	5.00	5.00	5.00
13		Yes	4.67	4.33	5.00
14		Yes	4.00	5.00	5.00
15		Yes	4.00	5.00	5.00
16		Yes	1.67	1.33	5.00
17		Yes	4.00	5.00	5.00
18		Yes	5.00	3.67	5.00
19		No	-	-	5.00
20		Yes	5.00	5.00	5.00
Average (Percentage)		**90 %**	**4.13 (82.6 %)**	**4.48 (89.6 %)**	**4.77 (95.4 %)**

contract, while GPT-4 + RAG grounds its response in the original document context to accurately extract the price adjustment formula. This side-by-side comparison of real queries demonstrates how RAG augmentation improves faithfulness by reducing hallucination and retrieving corroborating evidence to support generative outputs. As evidenced by the screenshots in Fig. 10, the GPT-4 + RAG model was deployed using the Streamlit application.

5.5. Model limitation

While the retrieval-augmented GPT-4 model shows promising results in querying the construction contract document, some limitations need to be acknowledged. The model struggled to retrieve relevant passages for two of the questions, defaulting to uninformative "I don't know" non-answers. This indicates that the chunking strategies and semantic search techniques used were unable to adequately link some complex questions to supporting evidence in the contract document. Future studies can explore different embedding models and vector databases and compare their performance. Generalization is another limitation - the model was trained on just a single contract document and may fail to transfer to new projects with different terminology, formats, and content. Training on a large corpus of diverse contracts would likely improve out-of-domain robustness.

6. Conclusion

This research aimed to comprehensively analyze the current state, opportunities, and challenges of applying GenAI in the construction industry. Key insights were synthesized through a systematic literature review and expert Delphi study. The literature review revealed that GenAI adoption in construction is still in very early stages, with just a handful of initial studies exploring applications like information retrieval, project planning, hazard recognition, and risk assessment. However, the great potential of generative techniques like LLMs was highlighted for enhancing productivity, accuracy, and automation across construction tasks.

The expert panel discussions further expanded on promising applications of GenAI in the construction industry during the pre-construction, construction, and post-construction phases. Opportunities were identified for major data types, including text, images, and video, with 76 potential applications mapped across different construction phases. The study also identified 18 critical challenges distributed across domain-specific, technological, adoption, and ethical categories. These challenges mirror those reported in other industries adopting GenAI but with construction-specific nuances. For instance, while data privacy represents a universal concern across sectors, the construction industry faces unique sensitivities regarding proprietary specifications and bid documentation, stemming from the competitive nature of construction bidding and the significant intellectual property

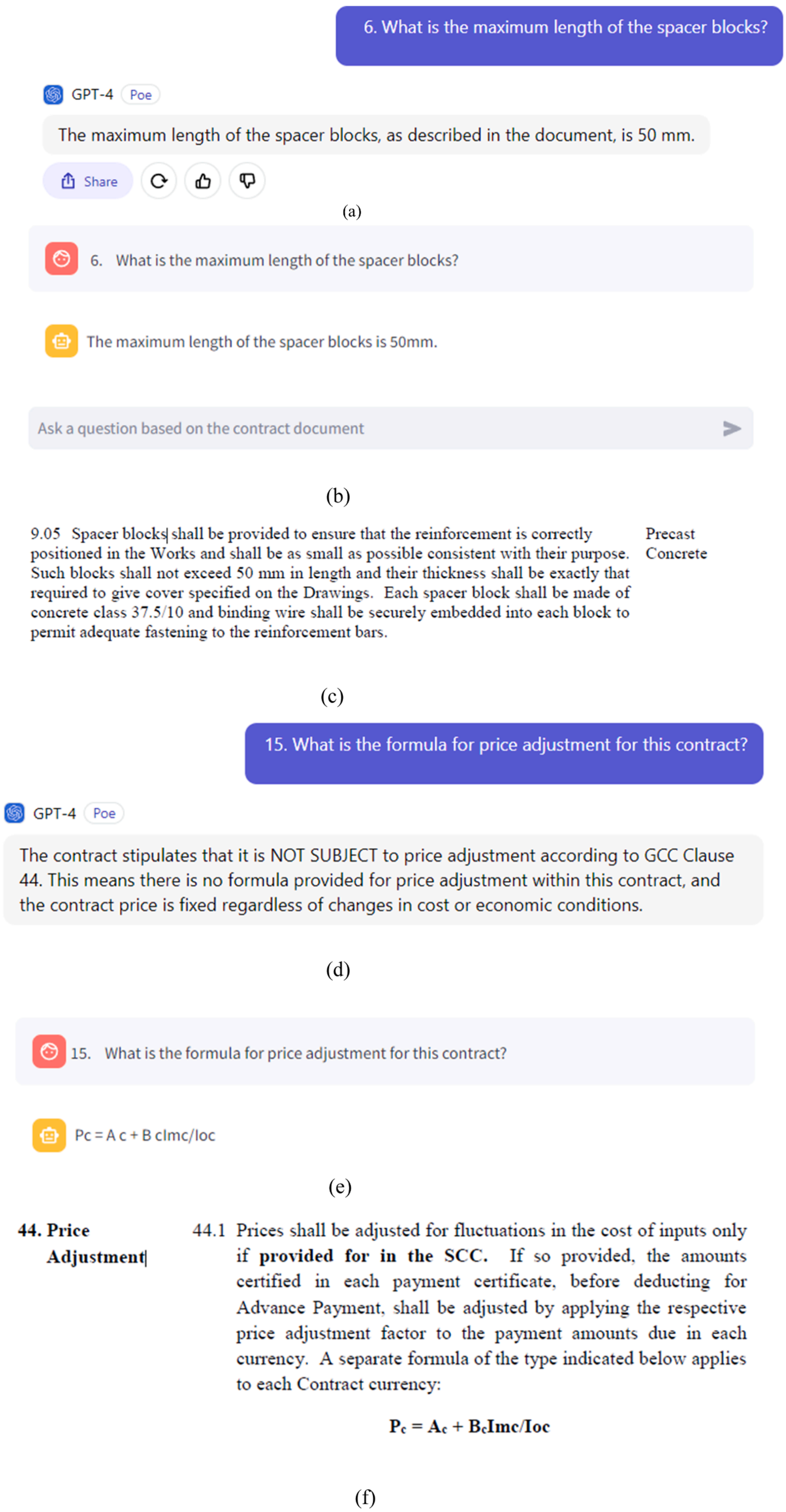


Fig. 10. Screenshots of queries, responses, and original text from the contract document.

embedded in design and construction methods.

A methodology was proposed to guide construction professionals in building customized GenAI solutions using their own proprietary data. The framework steps of data collection, preprocessing, model development, evaluation, and deployment aim to make these powerful technologies more accessible for practical industry deployment. The value of the framework was demonstrated through a case study on applying generative models for enhanced querying of construction contract documents. The retrieval-augmented system (RAG) showed a significantly improved ability to extract accurate, relevant information from contracts through natural language queries compared to a baseline generative model (GPT-4). In terms of quality, relevance, and reproducibility, the RAG system outperforms the base GPT-4 model by 5.2, 9.4, and 4.8 %, respectively.

While this study provides valuable insights into the application of GenAI in construction, certain limitations present opportunities for future work. The systematic literature review was confined to three databases - Scopus, Web of Science, and ScienceDirect - which yielded only 79 initial results, narrowing to 6 relevant papers after screening. Despite snowball searching, this limited scope may have missed relevant articles. The Delphi study involved 11 experts participating in three rounds, with a 73 % response rate from the 15 initially invited experts. While these experts had significant experience (81 % with over 10 years), and diverse backgrounds (27 % each in civil engineering, computer science, and construction management), the panel size was restricted due to resource constraints. For the case study, only a single base LLM (GPT-4) and embedding technique were utilized due to API access costs, with testing limited to 20 expert-validated queries on one construction contract. Testing multiple state-of-the-art models with RAG could reveal further performance gains beyond the current improvements of 5.2 %, 9.4 %, and 4.8 % in quality, relevance, and reproducibility, respectively. Future studies can build on these findings by expanding the literature review across more databases, recruiting larger expert panels as GenAI becomes more prominent in the construction industry, and experimenting with diverse generative architectures and embedding methods given sufficient computing power and budgets. Addressing these limitations represents an avenue for additional research and comparative assessment on applying GenAI for construction tasks. Nonetheless, this study provides a solid foundation of insights and a practical framework to guide further advancement.

CRediT authorship contribution statement

Ridwan Taiwo: Writing – original draft, Writing – review & editing, Visualization, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Idris Temitope Bello:** Writing – original draft, Writing – review & editing, Methodology, Visualization, Formal analysis. **Sulemana Fatoama Abdulai:** Writing – original draft, Writing – review & editing, Methodology, Investigation, Formal analysis, Data curation. **Abdul-Mugis Yussif:** Writing – review & editing, Methodology, Investigation. **Babatunde Abiodun Salami:** Writing – review & editing, Methodology, Visualization. **Abdullahi Saka:** Writing – review & editing, Methodology, Conceptualization. **Mohamed El Amine Ben Seghier:** Writing – review & editing, Supervision. **Tarek Zayed:** Writing – review & editing, Supervision, Project administration, Methodology.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors gratefully acknowledge the support from the Innovation and Technology Fund (Innovation and Technology Support Programme (ITSP)) under grant number ITS/033/20FP, and the Department of Building and Real Estate, the Hong Kong Polytechnic University. The corresponding author acknowledges funding from the European Union's Horizon 2021 research and innovation programme under the Marie Sklodowska-Curie project No- 101061320 and the funded APC by Oslo Metropolitan University.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.aej.2024.12.079.

References

[1] F. Barbosa, J. Woetzel, J. Mischke, F. Barbosa, M. Sridhar, M.J. Ribeirinho, M. Parsons, N. Bertram, S. Brown, Reinventing construction: a route to higher productivity, McKinsey Global Institute, 2017, pp. 7–13.
[2] J. Wells, The construction industry in the context of development: a new perspective, Habitat Int 8 (1984) 9–28, https://doi.org/10.1016/0197-3975(84)90040-7.
[3] M.Á. García, Challenges of the construction sector in the global economy and the knowledge society, Int. J. Strateg. Prop. Manag. 9 (2005) 65–77, https://doi.org/10.1080/1648715X.2005.9637528.
[4] MyJobQuote, The Latest Construction Industry Statistics, MyJobQuote, 2024 https://www.myjobquote.co.uk/blog/the-latest-construction-industry-statistics (accessed January 22, 2024).
[5] M. García, Moulding the future of construction, Concr. Eng. Int. 8 (2004) 41–42, https://www.scopus.com/inward/record.uri?eid= 2-s2.0-5144225594&partnerID= 40&md5= 082b9e7ffc41c7ca4414083e22be5c21.
[6] Edkins, A. Managing the design process in construction: A cognitive approach (1998). (Doctoral thesis, University of London). UCL Discovery. https://discovery.ucl.ac.uk/id/eprint/32848.
[7] T.M. Toole, M. Hallowell, Building performance engineering during construction, Constr. Res. Congr. 2005: Broadening Perspect. - Proc. Congr. (2005) 1–11, https://doi.org/10.1061/40754(183)7.
[8] Wolfe Jr, S., 2020 Report: Construction Suffers From Wasted Time & Slow Payment. Levelset (2020). Retrieved from https://www.levelset.com/blog/2020-report-construction-wasted-time-slow-payment/.
[9] Boussabaine, A., & Kirkham, R. Whole Life-Cycle Costing: Risk and Risk Responses. John Wiley & Sons, (2008). Retrieved from Google Books.
[10] T.M. Toole, M. Hallowell, Building Performance Engineering during Construction, Construction Research Congress 2005: Broadening Perspectives - Proceedings of the Congress (2005) 1–11. https://doi.org/10.1061/40754(183)7.
[11] J. Lee, D. Jung, J. Moon, S. Rho, Advanced R-GAN: Generating anomaly data for improved detection in imbalanced datasets using regularized generative adversarial networks, Alex. Eng. J. 111 (2025) 491–510, https://doi.org/10.1016/j.aej.2024.10.084.
[12] A. Taparia, A.K. Bashir, Y. Zhu, T.R. Gdekallu, K. Nath, Transforming satellite imagery into vector maps using modified GANs, Alex. Eng. J. 109 (2024) 792–806, https://doi.org/10.1016/j.aej.2024.09.074.
[13] O. Nagy, I. Papp, R.Z. Szabó, Construction 4.0 organisational level challenges and solutions, Sustainability 13 (2021) 12321, https://doi.org/10.3390/SU132112321.
[14] Z. Chen, Grand challenges in construction management, Front Built Environ. 5 (2019) 415116, https://doi.org/10.3389/FBUIL.2019.00031/BIBTEX.
[15] M. Parsamehr, U.S. Perera, T.C. Dodanwala, P. Perera, R. Ruparathna, A review of construction management challenges and BIM-based solutions: perspectives from the schedule, cost, quality, and safety management, Asian J. Civ. Eng. 24 (2022) 353–389, https://doi.org/10.1007/S42107-022-00501-4.
[16] M. Bilal, L.O. Oyedele, J. Qadir, K. Munir, S.O. Ajayi, O.O. Akinade, H. A. Owolabi, H.A. Alaka, M. Pasha, Big Data in the construction industry: a review of present status, opportunities, and future trends, Adv. Eng. Inform. 30 (2016) 500–521, https://doi.org/10.1016/j.aei.2016.07.001.
[17] A.B. Saka, L.O. Oyedele, L.A. Akanbi, S.A. Ganiyu, D.W.M. Chan, S.A. Bello, Conversational artificial intelligence in the AEC industry: a review of present status, challenges and opportunities, Adv. Eng. Inform. 55 (2023) 101869, https://doi.org/10.1016/j.aei.2022.101869.
[18] S. Sachan, X. Liu (Lisa), Blockchain-based auditing of legal decisions supported by explainable AI and generative AI tools, Eng. Appl. Artif. Intell. 129 (2024), https://doi.org/10.1016/j.engappai.2023.107666.
[19] B. Regassa Hunde, A. Debebe Woldeyohannes, Future prospects of computer-aided design (CAD) – a review from the perspective of artificial intelligence (AI), extended reality, and 3D printing, Results Eng. 14 (2022), https://doi.org/10.1016/j.rineng.2022.100478.
[20] A.M. Rinaldi, C. Russo, C. Tommasino, Automatic image captioning combining natural language processing and deep neural networks, Results Eng. 18 (2023), https://doi.org/10.1016/j.rineng.2023.101107.
[21] A. Alexiadis, B. Ghiassi, From text to tech: shaping the future of physics-based simulations with AI-driven generative models, Results Eng. 21 (2024), https://doi.org/10.1016/j.rineng.2023.101721.

[22] J. Yang, H. Jin, R. Tang, X. Han, Q. Feng, H. Jiang, B. Yin, X. Hu, Harness Power LLMs Pract. A Surv. ChatGPT Beyond 1 (2023) 1–24.
[23] J. Pang, Y. Wang, F. Ozyurt, S. Dogan, T. Tuncer, L. Yu, Utilizing language models for advanced electrocardiogram analysis, Alex. Eng. J. 105 (2024) 460–470, https://doi.org/10.1016/j.aej.2024.07.086.
[24] P. Yang, S. Li, S. Qin, L. Wang, M. Hu, F. Yang, Smart grid enterprise decision-making and economic benefit analysis based on LSTM-GAN and edge computing algorithm, Alex. Eng. J. 104 (2024) 314–327, https://doi.org/10.1016/j.aej.2024.06.028.
[25] S.A. Prieto, E.T. Mengiste, B. García de Soto, Investigating the use of ChatGPT for the scheduling of construction projects, Buildings 13 (4) (2023) 857, https://doi.org/10.3390/buildings13040857.
[26] A. Saka, R. Taiwo, N. Saka, B.A. Salami, S. Ajayi, K. Akande, H. Kazemi, GPT models in construction industry: opportunities, limitations, and a use case validation, Dev. Built Environ. 17 (2024) 100300, https://doi.org/10.1016/j.dibe.2023.100300.
[27] OpenAI, DALL·E 3, OpenAI (2023). https://openai.com/dall-e-3 (accessed December 25, 2023).
[28] Google, Imagen Video, Google (2022). https://imagen.research.google/video/ (accessed January 25, 2024).
[29] A. You, J.K. Kim, I.H. Ryu, T.K. Yoo, Application of generative adversarial networks (GAN) for ophthalmology image domains: a survey, Eye Vis. 9 (2022), https://doi.org/10.1186/s40662-022-00277-3.
[30] R. Gómez-Bombarelli, J.N. Wei, D. Duvenaud, J.M. Hernández-Lobato, B. Sánchez-Lengeling, D. Sheberla, J. Aguilera-Iparraguirre, T.D. Hirzel, R. P. Adams, A. Aspuru-Guzik, Automatic chemical design using a data-driven continuous representation of molecules, ACS Cent. Sci. 4 (2018) 268–276, https://doi.org/10.1021/acscentsci.7b00572.
[31] I.T. Bello, R. Taiwo, O.C. Esan, A.H. Adegoke, A.O. Ijaola, Z. Li, S. Zhao, C. Wang, Z. Shao, M. Ni, AI-enabled materials discovery for advanced ceramic electrochemical cells, Energy AI 15 (2024) 100317, https://doi.org/10.1016/J.EGYAI.2023.100317.
[32] M.X. Tang, J. Cui, Towards generative systems for supporting product design, Int. J. Des. Eng. 7 (1) (2017), https://doi.org/10.1504/ijde.2017.10006425.
[33] C.A. Bali, Can Generative AI improve social science? Proc. Natl. Acad. Sci. 121 (21) (2024) e2314021121 https://doi.org/10.1073/pnas.2314021121.
[34] R. Kokku, S. Sundararajan, P. Dey, R. Sindhgatta, S. Nitta, B. Sengupta, Augmenting Classrooms with AI for Personalized Education, ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings 2018-April (2018) 6976–6980. https://doi.org/10.1109/ICASSP.2018.8461812.
[35] P. Danilo, Introducing Llama 3.2 models from Meta in Amazon Bedrock: A new generation of multimodal vision and lightweight models, (2024).
[36] B. Yask, Grok-1: A Massive 314 Billion Parameter Language Model Released Open Source, 2024.
[37] Flux Pro Art, What is Flux Dev?, (2024). Retrieved from https://fluxpro.art/posts/about.
[38] Anthropic, The Claude 3 Model Family: Opus, Sonnet, Haiku Anthropic, (2024). Retrieved from https://docs.anthropic.com/.
[39] R. Alberto, OpenAI o1: A New Paradigm For AI, (2024).
[40] C. Erfei, H. Yinan, M. Zheng, T. Hao, W. Weiyun, L. Kunchang, W. Yi, W. Wenhai, Z. Xizhou, L. Lewei, L. Tong, W. Yali, W. Limin, Q. Yu, D. Jifeng, Comprehensive Multimodal Annotations With GPT-4o, (2024).
[41] OpenAI, :, J. Achiam, S. Adler, S. Agarwal, L. Ahmad, I. Akkaya, F.L. Aleman, D. Almeida, J. Altenschmidt, S. Altman, S. Anadkat, R. Avila, I. Babuschkin, S. Balaji, V. Balcom, P. Baltescu, H. Bao, M. Bavarian, J. Belgum, I. Bello, J. Berdine, G. Bernadett-Shapiro, C. Berner, L. Bogdonoff, O. Boiko, M. Boyd, A.-L. Brakman, G. Brockman, T. Brooks, M. Brundage, K. Button, T. Cai, R. Campbell, A. Cann, B. Carey, C. Carlson, R. Carmichael, B. Chan, C. Chang, F. Chantzis, D. Chen, S. Chen, R. Chen, J. Chen, M. Chen, B. Chess, C. Cho, C. Chu, H.W. Chung, D. Cummings, J. Currier, Y. Dai, C. Decareaux, T. Degry, N. Deutsch, D. Deville, A. Dhar, D. Dohan, S. Dowling, S. Dunning, A. Ecoffet, A. Eleti, T. Eloundou, D. Farhi, L. Fedus, N. Felix, S.P. Fishman, J. Forte, I. Fulford, L. Gao, E. Georges, C. Gibson, V. Goel, T. Gogineni, G. Goh, R. Gontijo-Lopes, J. Gordon, M. Grafstein, S. Gray, R. Greene, J. Gross, S.S. Gu, Y. Guo, C. Hallacy, J. Han, J. Harris, Y. He, M. Heaton, J. Heidecke, C. Hesse, A. Hickey, W. Hickey, P. Hoeschele, B. Houghton, K. Hsu, S. Hu, X. Hu, J. Huizinga, S. Jain, S. Jain, J. Jang, A. Jiang, R. Jiang, H. Jin, D. Jin, S. Jomoto, B. Jonn, H. Jun, T. Kaftan, Ł. Kaiser, A. Kamali, I. Kanitscheider, N.S. Keskar, T. Khan, L. Kilpatrick, J.W. Kim, C. Kim, Y. Kim, H. Kirchner, J. Kiros, M. Knight, D. Kokotajlo, Ł. Kondraciuk, A. Kondrich, A. Konstantinidis, K. Kosic, G. Krueger, V. Kuo, M. Lampe, I. Lan, T. Lee, J. Leike, J. Leung, D. Levy, C.M. Li, R. Lim, M. Lin, S. Lin, M. Litwin, T. Lopez, R. Lowe, P. Lue, A. Makanju, K. Malfacini, S. Manning, T. Markov, Y. Markovski, B. Martin, K. Mayer, A. Mayne, B. McGrew, S.M. McKinney, C. McLeavey, P. McMillan, J. McNeil, D. Medina, A. Mehta, J. Menick, L. Metz, A. Mishchenko, P. Mishkin, V. Monaco, E. Morikawa, D. Mossing, T. Mu, M. Murati, O. Murk, D. Mély, A. Nair, R. Nakano, R. Nayak, A. Neelakantan, R. Ngo, H. Noh, L. Ouyang, C. O'Keefe, J. Pachocki, A. Paino, J. Palermo, A. Pantuliano, G. Parascandolo, J. Parish, E. Parparita, A. Passos, M. Pavlov, A. Peng, A. Perelman, F. de A.B. Peres, M. Petrov, H.P. de O. Pinto, Michael, Pokorny, M. Pokrass, V. Pong, T. Powell, A. Power, B. Power, E. Proehl, R. Puri, A. Radford, J. Rae, A. Ramesh, C. Raymond, F. Real, K. Rimbach, C. Ross, B. Rotsted, H. Roussez, N. Ryder, M. Saltarelli, T. Sanders, S. Santurkar, G. Sastry, H. Schmidt, D. Schnurr, J. Schulman, D. Selsam, K. Sheppard, T. Sherbakov, J. Shieh, S. Shoker, P. Shyam, S. Sidor, E. Sigler, M. Simens, J. Sitkin, K. Slama, I. Sohl, B. Sokolowsky, Y. Song, N. Staudacher, F.P. Such, N. Summers, I. Sutskever, J. Tang, N. Tezak, M. Thompson, P. Tillet, A. Tootoonchian, E. Tseng, P. Tuggle, N. Turley, J. Tworek, J.F.C. Uribe, A. Vallone, A. Vijayvergiya, C. Voss, C. Wainwright, J.J. Wang, A. Wang, B. Wang, J. Ward, J. Wei, C. Weinmann, A. Welihinda, P. Welinder, J. Weng, L. Weng, M. Wiethoff, D. Willner, C. Winter, S. Wolrich, H. Wong, L. Workman, S. Wu, J. Wu, M. Wu, K. Xiao, T. Xu, S. Yoo, K. Yu, Q. Yuan, W. Zaremba, R. Zellers, C. Zhang, M. Zhang, S. Zhao, T. Zheng, J. Zhuang, W. Zhuk, B. Zoph, G.P.T.-4 Technical report, (2023). ⟨http://arxiv.org/abs/2303.08774⟩.
[42] S. Pichai, D. Hassabis, Introducing Gemini: our largest and most capable AI model, (2023). https://blog.google/technology/ai/google-gemini-ai/ (accessed February 9, 2024).
[43] Schmid, P. (2024). LLaMA 2 - Every Resource you need. Retrieved from https://philschmid.de/llama-2.
[44] A. Chowdhery, S. Narang, J. Devlin, M. Bosma, G. Mishra, A. Roberts, P. Barham, H.W. Chung, C. Sutton, S. Gehrmann, P. Schuh, K. Shi, S. Tsvyashchenko, J. Maynez, A. Rao, P. Barnes, Y. Tay, N. Shazeer, V. Prabhakaran, E. Reif, N. Du, B. Hutchinson, R. Pope, J. Bradbury, J. Austin, M. Isard, G. Gur-Ari, P. Yin, T. Duke, A. Levskaya, S. Ghemawat, S. Dev, H. Michalewski, X. Garcia, V. Misra, K. Robinson, L. Fedus, D. Zhou, D. Ippolito, D. Luan, H. Lim, B. Zoph, A. Spiridonov, R. Sepassi, D. Dohan, S. Agrawal, M. Omernick, A.M. Dai, T.S. Pillai, M. Pellat, A. Lewkowycz, E. Moreira, R. Child, O. Polozov, K. Lee, Z. Zhou, X. Wang, B. Saeta, M. Diaz, O. Firat, M. Catasta, J. Wei, K. Meier-Hellstern, D. Eck, J. Dean, S. Petrov, N. Fiedel, PaLM: Scaling Language Modeling with Pathways, (2022). ⟨http://arxiv.org/abs/2204.02311⟩.
[45] D. Podell, Z. English, K. Lacey, A. Blattmann, T. Dockhorn, J. Müller, J. Penna, R. Rombach, SDXL Improv. Latent Diffus. Models High. Resolut. Image Synth. (2023). ⟨http://arxiv.org/abs/2307.01952⟩.
[46] V. Thengane, S. Khan, M. Hayat, F. Khan, CLIP Model Is. Effic. Contin. Learn. (2022). ⟨http://arxiv.org/abs/2210.03114⟩.
[47] B. Poole, A. Jain, J.T. Barron, B. Mildenhall, DreamFusion Text. to-3D Using 2D Diffus. (2022). ⟨http://arxiv.org/abs/2209.14988⟩.
[48] J.-B. Alayrac, J. Donahue, P. Luc, A. Miech, I. Barr, Y. Hasson, K. Lenc, A. Mensch, K. Millican, M. Reynolds, R. Ring, E. Rutherford, S. Cabi, T. Han, Z. Gong, S. Samangooei, M. Monteiro, J. Menick, S. Borgeaud, A. Brock, A. Nematzadeh, S. Sharifzadeh, M. Binkowski, R. Barreira, O. Vinyals, A. Zisserman, K. Simonyan, Flamingo. a Vis. Lang. Model Few-Shot Learn. (2022). ⟨http://arxiv.org/abs/2204.14198⟩.
[49] R. Villegas, M. Babaeizadeh, P.-J. Kindermans, H. Moraldo, H. Zhang, M.T. Saffar, S. Castro, J. Kunze, D. Erhan, Phenaki Var. Length Video Gener. Open Domain Textual Descr. (2022). ⟨http://arxiv.org/abs/2210.02399⟩.
[50] M. Chen, J. Tworek, H. Jun, Q. Yuan, H.P. de O. Pinto, J. Kaplan, H. Edwards, Y. Burda, N. Joseph, G. Brockman, A. Ray, R. Puri, G. Krueger, M. Petrov, H. Khlaaf, G. Sastry, P. Mishkin, B. Chan, S. Gray, N. Ryder, M. Pavlov, A. Power, L. Kaiser, M. Bavarian, C. Winter, P. Tillet, F.P. Such, D. Cummings, M. Plappert, F. Chantzis, E. Barnes, A. Herbert-Voss, W.H. Guss, A. Nichol, A. Paino, N. Tezak, J. Tang, I. Babuschkin, S. Balaji, S. Jain, W. Saunders, C. Hesse, A.N. Carr, J. Leike, J. Achiam, V. Misra, E. Morikawa, A. Radford, M. Knight, M. Brundage, M. Murati, K. Mayer, P. Welinder, B. McGrew, D. Amodei, S. McCandlish, I. Sutskever, W. Zaremba, Evaluating Large Language Models Trained on Code, (2021). ⟨http://arxiv.org/abs/2107.03374⟩.
[51] R. Taylor, M. Kardas, G. Cucurull, T. Scialom, A. Hartshorn, E. Saravia, A. Poulton, V. Kerkez, R. Stojnic, Galactica: A Large Language Model for Science, (2022). http://arxiv.org/abs/2211.09085.
[52] Z. Yang, Z. Dai, Y. Yang, J. Carbonell, R. Salakhutdinov, Q.V. Le, XLNet: Generalized Autoregressive Pretraining for Language Understanding, (2019). ⟨http://arxiv.org/abs/1906.08237⟩.
[53] K.S. Michael, What is generative AI? Everything you need to know, (2023). https://www.techtarget.com/searchenterpriseai/definition/generative-AI (accessed February 9, 2024).
[54] M. Lewis, Y. Liu, N. Goyal, M. Ghazvininejad, A. Mohamed, O. Levy, V. Stoyanov, L. Zettlemoyer, BART: Denoising Sequence-to-Sequence Pre-training for Natural Language Generation, Translation, and Comprehension, (2019). http://arxiv.org/abs/1910.13461.
[55] C. Raffel, N. Shazeer, A. Roberts, K. Lee, S. Narang, M. Matena, Y. Zhou, W. Li, P. J. Liu, Explor. Limits Transf. Learn. a Unifi-.-. Text. to Text. Transform. (2019). ⟨http://arxiv.org/abs/1910.10683⟩.
[56] A. Mandour, GPT-3.5 model architecture, (2022). https://iq.opengenus.org/gpt-3-5-model/ (accessed February 8, 2024).
[57] A. Radford, J. Wu, R. Child, D. Luan, D. Amodei, I. Sutskever, Language Models are Unsupervised Multitask Learners, (2022). https://github.com/codelucas/newspaper.
[58] J. Zheng, M. Fischer, Dynamic prompt-based virtual assistant framework for BIM information search, Autom. Constr. 155 (2023) 105067, https://doi.org/10.1016/j.autcon.2023.105067.
[59] W. Liao, X. Lu, Y. Fei, Y. Gu, Y. Huang, Generative AI design for building structures, Autom. Constr. 157 (2024), https://doi.org/10.1016/j.autcon.2023.105187.
[60] P. Ghimire, K. Kim, Opportunities and Challenges of Generative AI in Construction Industry: Focusing on Adoption of Text-Based Models, Buildings 14 (2024) 220.
[61] R. Taiwo, I.A. Shaban, T. Zayed, Development of sustainable water infrastructure: a proper understanding of water pipe failure, J. Clean. Prod. 398 (2023).
[62] I.T. Bello, S. Zhai, S. Zhao, Z. Li, N. Yu, M. Ni, Scientometric review of proton-conducting solid oxide fuel cells, Int. J. Hydrog. Energy 46 (2021) 37406–37428, https://doi.org/10.1016/j.ijhydene.2021.09.061.

[63] Shaoni Mukherjee, LangChain: a beginner's guide to harness the power of language models, Paperspace (2024) https://blog.paperspace.com/langchain/ (accessed February 8, 2024).
[64] N. Mhadbi, Python Tutorial: Streamlit, Datacamp (2021). https://www.datacamp.com/tutorial/streamlit (accessed February 2, 2024).
[65] S.M.J. Uddin, A. Albert, A. Ovid, A. Alsharef, Leveraging ChatGPT to aid construction hazard recognition and support safety education and training, Sustainability 15 (2023), https://doi.org/10.3390/su15097121.
[66] F. Amer, Y. Jung, M. Golparvar-Fard, Transformer machine learning language model for auto-alignment of long-term and short-term plans in construction, Autom. Constr. 132 (2021) 103929, https://doi.org/10.1016/j.autcon.2021.103929.
[67] S. Moon, S. Chi, S.B. Im, Automated detection of contractual risk clauses from construction specifications using bidirectional encoder representations from transformers (BERT), Autom. Constr. 142 (2022) 104465, https://doi.org/10.1016/j.autcon.2022.104465.
[68] H. You, Y. Ye, T. Zhou, Q. Zhu, J. Du, Robot-enabled construction assembly with automated sequence planning based on ChatGPT: RoboGPT, Buildings 13 (2023), https://doi.org/10.3390/buildings13071772.
[69] S.M.J. Uddin, A. Albert, A. Ovid, A. Alsharef, Leveraging ChatGPT to AId Construction Hazard Recognition and Support Safety Education and Training, Sustainability 15 (2023), https://doi.org/10.3390/su15097121.
[70] S. Moon, S. Chi, S.B. Im, Automated detection of contractual risk clauses from construction specifications using bidirectional encoder representations from transformers (BERT), Autom. Constr. 142 (2022) 104465, https://doi.org/10.1016/j.autcon.2022.104465.
[71] F. Amer, Y. Jung, M. Golparvar-Fard, Transformer machine learning language model for auto-alignment of long-term and short-term plans in construction, Autom. Constr. 132 (2021) 103929, https://doi.org/10.1016/j.autcon.2021.103929.
[72] OpenAI, GPT-4, OpenAI (2023). https://openai.com/research/gpt-4 (accessed January 25, 2023).
[73] DeepMind, Welcome to the Gemini era, Google (2023). https://deepmind.google/technologies/gemini/#introduction (accessed January 25, 2024).
[74] W. Hong, M. Ding, W. Zheng, X. Liu, J. Tang, CogVideo: large-scale pretraining for text-to-video generation via transformers, ArXiv Prepr. (2022) 1–15.
[75] O. Bar-Tal, H. Chefer, O. Tov, C. Herrmann, R. Paiss, S. Zada, A. Ephrat, J. Hur, Y. Li, T. Michaeli, O. Wang, D. Sun, T. Dekel, I. Mosseri, Lumiere: A Space-Time Diffusion Model for Video Generation, (2024).
[76] P. Esser, J. Chiu, P. Atighehchian, J. Granskog, A. Germanidis, Structure and Content-Guided Video Synthesis with Diffusion Models, (2023).
[77] P. Isola, J.Y. Zhu, T. Zhou, A.A. Efros, Image-to-image translation with conditional adversarial networks, ArXiv Prepr. (2018), https://doi.org/10.1109/CVPR.2017.632.
[78] A. Blattmann, T. Dockhorn, S. Kulal, D. Mendelevitch, M. Kilian, D. Lorenz, Y. Levi, Z. English, V. Voleti, A. Letts, V. Jampani, R. Rombach, Stable Video Diffusion: Scaling Latent Video Diffusion Models to Large Datasets, (2023). http://arxiv.org/abs/2311.15127 (accessed January 25, 2024).
[79] S. Yan, T. Zhu, Z. Wang, Y. Cao, M. Zhang, S. Ghosh, Y. Wu, J. Yu, VideoCoCa: video-text modeling with zero-shot, Transf. Contrastive Captioners (2022).
[80] J. Dimyadi, P. Pauwels, M. Spearpoint, C. Clifton, R. Amor, Querying a Regulatory Model for Compliant Building Design Audit, Proc. of the 32nd CIB W78 Conference 2015, 27th-29th October 2015, Eindhoven, The Netherlands (2015) 139–148. https://doi.org/10.13140/RG.2.1.4022.6003.
[81] E.P. Karan, J. Irizarry, Extending BIM interoperability to preconstruction operations using geospatial analyses and semantic web services, Autom. Constr. 53 (2015) 1–12, https://doi.org/10.1016/j.autcon.2015.02.012.
[82] R. Edirisinghe, H. Lingard, Exploring the potential for the use of video to communicate safety information to construction workers: case studies of organizational use*, Constr. Manag. Econ. 34 (2016) 366–376, https://doi.org/10.1080/01446193.2016.1200736.
[83] J. Yang, M.W. Park, P.A. Vela, M. Golparvar-Fard, Construction performance monitoring via still images, time-lapse photos, and video streams: now, tomorrow, and the future, Adv. Eng. Inform. 29 (2015) 211–224, https://doi.org/10.1016/j.aei.2015.01.011.
[84] T. Rakha, A. Gorodetsky, Review of unmanned aerial system (UAS) applications in the built environment: towards automated building inspection procedures using drones, Autom. Constr. 93 (2018) 252–264, https://doi.org/10.1016/j.autcon.2018.05.002.
[85] Z. Chen, K. Yi, Y. Li, M. Ding, A. Torralba, J.B. Tenenbaum, C. Gan, Comphy: Compositional Physical Reasoning of Objects and Events From Videos, ICLR 2022 - 10th International Conference on Learning Representations (2022) 1–18.
[86] L.S. Kang, H.S. Kim, H.S. Moon, S.K. Kim, Managing construction schedule by telepresence: integration of site video feed with an active nD CAD simulation, Autom. Constr. 68 (2016) 32–43, https://doi.org/10.1016/j.autcon.2016.04.003.
[87] Y. Zhang, X. Liang, D. Zhang, M. Tan, E.P. Xing, Unsupervised object-level video summarization with online motion auto-encoder, Pattern Recognit. Lett. 130 (2020) 376–385, https://doi.org/10.1016/j.patrec.2018.07.030.
[88] F.H. Amr, M.A. Mona, M.O. Hesham, Optimizing labor productivity in Egypt using regression prediction models, Proc., Annu. Conf. Can. Soc. Civ. Eng. (2019).
[89] O. Okudan, C. Budayan, I. Dikmen, A knowledge-based risk management tool for construction projects using case-based reasoning, Expert Syst. Appl. 173 (2021) 114776, https://doi.org/10.1016/j.eswa.2021.114776.
[90] C.J. Liang, T.H. Le, Y. Ham, B.R.K. Mantha, M.H. Cheng, J.J. Lin, Ethics of artificial intelligence and robotics in the architecture, engineering, and construction industry, Autom. Constr. 162 (2024), https://doi.org/10.1016/j.autcon.2024.105369.
[91] S. Jang, G. Lee, Interactive design by integrating a large pre-trained language model and building information modeling suhyung, ArXiv Prepr. (2022).
[92] T.O. Ayodele, K. Kajimo-Shakantu, The fourth industrial revolution (4thIR) and the construction industry - the role of data sharing and assemblage, IOP Conf. Ser. Earth Environ. Sci. 654 (2021), https://doi.org/10.1088/1755-1315/654/1/012013.
[93] P. Pauwels, T.M. De Farias, C. Zhang, A. Roxin, J. Beetz, J. De Roo, C. Nicolle, Querying and reasoning over large scale building data sets: An outline of a performance benchmark, Proceedings of the ACM SIGMOD International Conference on Management of Data (2016). https://doi.org/10.1145/2928294.2928303.
[94] H.C. Yeung, T. Ridwan, S. Tariq, T. Zayed, BEAM Plus implementation in Hong Kong: assessment of challenges and policies, Int. J. Constr. Manag. 0 (2020) 1–15, https://doi.org/10.1080/15623599.2020.1827692.
[95] A. Kumar, G. Ghinea, S. Merugu, Proceedings of the 2nd International Conference on Cognitive and Intelligent Computing, 2022.
[96] B. Franci, S. Grammatico, Training generative adversarial networks via stochastic nash games, IEEE Trans. Neural Netw. Learn Syst. 34 (2023) 1319–1328, https://doi.org/10.1109/TNNLS.2021.3105227.
[97] R. Taiwo, T. Zayed, M.E.A. Ben Seghier, Integrated intelligent models for predicting water pipe failure probability, Alex. Eng. J. 86 (2024) 243–257, https://doi.org/10.1016/j.aej.2023.11.047.
[98] Z. Yang, L. Meng, J.-W. Chung, Chasing Low-Carbon Electricity for Practical and Sustainable DNN Training. Tackling Climate Change with Machine Learning, ICLR 2023 (2023). https://arxiv.org/pdf/2303.02508.
[99] D. Ganguli, D. Hernandez, L. Lovitt, A. Askell, Y. Bai, A. Chen, T. Conerly, N. Dassarma, D. Drain, N. Elhage, S. El Showk, S. Fort, Z. Hatfield-Dodds, T. Henighan, S. Johnston, A. Jones, N. Joseph, J. Kernian, S. Kravec, B. Mann, N. Nanda, K. Ndousse, C. Olsson, D. Amodei, T. Brown, J. Kaplan, S. McCandlish, C. Olah, D. Amodei, J. Clark, Predictability and Surprise in Large Generative Models, ACM International Conference Proceeding Series (2022) 1747–1764. https://doi.org/10.1145/3531146.3533229.
[100] A.N. Wu, R. Stouffs, F. Biljecki, Generative Adversarial Networks in the built environment: a comprehensive review of the application of GANs across data types and scales, Build. Environ. 223 (2022) 109477, https://doi.org/10.1016/j.buildenv.2022.109477.
[101] S. Na, S. Heo, S. Han, Y. Shin, Y. Roh, Acceptance model of artificial intelligence (AI)-based technologies in construction firms: applying the technology acceptance model (TAM) in combination with the technology–organisation–environment (TOE) framework, Buildings 12 (2022), https://doi.org/10.3390/buildings12020090.
[102] P.E.D. Love, W. Fang, J. Matthews, S. Porter, H. Luo, L. Ding, Explainable artificial intelligence (XAI): precepts, models, and opportunities for research in construction, Adv. Eng. Inform. 57 (2023) 1–58, https://doi.org/10.1016/j.aei.2023.102024.
[103] T.D. Akinosho, L.O. Oyedele, M. Bilal, A.O. Ajayi, M.D. Delgado, O.O. Akinade, A. A. Ahmed, Deep learning in the construction industry: a review of present status and future innovations, J. Build. Eng. 32 (2020) 101827, https://doi.org/10.1016/j.jobe.2020.101827.
[104] T. Cerovsek, A review and outlook for a "Building Information Model" (BIM): a multi-standpoint framework for technological development, Adv. Eng. Inform. 25 (2011) 224–244, https://doi.org/10.1016/j.aei.2010.06.003.
[105] P. Księżak, Toward a Conceptual Network for the Private Law of Artificial Intelligence, Springer (2023), https://doi.org/10.1007/978-3-031-19447-4.
[106] T. Patel, V. Patel, Data privacy in construction industry by privacy-preserving data mining (PPDM) approach, Asian J. Civ. Eng. 21 (2020) 505–515, https://doi.org/10.1007/s42107-020-00225-3.
[107] K. Maitz, A. Fessl, V. Pammer-Schindler, R. Kaiser, S. Lindstaedt, What do Construction Workers Know about Artificial Intelligence? An Exploratory Case Study in an Austrian SME, ACM International Conference Proceeding Series (2022) 389–393. https://doi.org/10.1145/3543758.3547545.
[108] W. Ma, Y. Song, M. Xue, S. Wen. The "Code" of Ethics: A Holistic Audit of, AI Code Generators. Journal of LaTeX Class Files, 2021.
[109] R. Taiwo, A. Yussif, M. El, A. Ben, T. Zayed, Explainable ensemble models for predicting wall thickness loss of water pipes, Ain Shams Eng. J. (2024) 102630, https://doi.org/10.1016/j.asej.2024.102630.
[110] T. Grønsund, M. Aanestad, Augmenting the algorithm: emerging human-in-the-loop work configurations, J. Strateg. Inf. Syst. 29 (2020) 101614, https://doi.org/10.1016/j.jsis.2020.101614.
[111] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A.N. Gomez, Ł. Kaiser, I. Polosukhin, Attention is all you need, Adv. Neural Inf. Process Syst. (2017) 5999–6009.
[112] R.K. Luu, M.J. Buehler, BioinspiredLLM: conversational Large Language Model for the Mechanics of Biological and Bio-Inspired Materials, Adv. Sci. (2023) 1–16, https://doi.org/10.1002/advs.202306724.
[113] M.Y. Cheng, H.C. Tsai, W.S. Hsieh, Web-based conceptual cost estimates for construction projects using Evolutionary Fuzzy Neural Inference Model, Autom. Constr. 18 (2009) 164–172, https://doi.org/10.1016/j.autcon.2008.07.001.
[114] O. Bentaleb, A.S.Z. Belloum, A. Sebaa, A. El-Maouhab, Containerization technologies: taxonomies, applications and challenges, J. Supercomput. 78 (2022) 1144–1181, https://doi.org/10.1007/s11227-021-03914-1.
[115] S. Shankar, R. Garcia, J. M. Hellerstein, & A. G. Parameswaran. Operationalizing Machine Learning: An Interview Study. arXiv. (2022). https://arxiv.org/2209.09125.

[116] K.S. Kalyan, A survey of GPT-3 family large language models including ChatGPT and GPT-4, Nat. Lang. Process. J. 6 (2024) 100048, https://doi.org/10.1016/j.nlp.2023.100048.
[117] Z. Jiang, F. Xu, L. Gao, Z. Sun, Q. Liu, J. Dwivedi-Yu, Y. Yang, J. Callan, G. Neubig, Active Retrieval Augmented Generation, (2023) 7969-7992. https://doi.org/10.18653/v1/2023.emnlp-main.495.
[118] A. Cassandra, Cassandra Documentation, Apache Cassandra (2023). https://cassandra.apache.org/doc/latest/cassandra/developing/cql/index.html (accessed February 7, 2024).
[119] M. Wölfel, M.B. Shirzad, A. Reich, K. Anderer, Knowledge-based and generative-ai-driven pedagogical conversational agents: a comparative study of grice's cooperative principles and trust, Big Data Cogn. Comput. 8 (2023) 2, https://doi.org/10.3390/bdcc8010002.