www.itcon.org - Journal of Information Technology in Construction - ISSN 1874-4753

ENHANCING REGULATORY COMPLIANCE IN AEC INDUSTRY VIA LLM-POWERED DECISION FRAMEWORKS

SUBMITTED: September 2025
PUBLISHED: February 2026
EDITOR: Bimal Kumar
DOI: 10.36680/j.itcon.2026.009

Entesar Al Nama, PhD
University of Bahrain
entesarjasim@gmail.com

Maqsood Mahmud, Assistant Professor
School of Computing, Faculty of Computing, Engineering and the Built Environment (CEBE), Ulster University, York Street, BT151AP, Belfast, United Kingdom
m.mahmud@ulster.ac.uk

Huda Al Madhoob, Dr.
College of Engineering, University of Bahrain
halmadhoob@uob.edu.bh

SUMMARY: *The AEC industry is a highly intricate ecosystem involving architects, structural engineers, civil consultants, and contractors working under intense time pressures, where the success of projects hinges on sound decision-making. One of the most persistent challenges is regulatory compliance—not merely understanding the rules, but accurately interpreting and applying them within real-world constraints. This research paper aims to enhance regulatory decision-making through AI, specifically a Q&A model, and stands out in five key ways: it tackles the widespread issue of building code violations that cause costly delays; it empowers professionals to verify compliance during the design phase, minimizing errors and saving resources; it proves that effective AI doesn't require massive datasets, instead leveraging domain expertise and smart data strategies; and it introduces a scalable framework that can extend to broader regulatory domains and integrate with BIM tools for automated checks, offering a transformative approach to compliance in the AEC sector.*

KEYWORDS: *Bahrain, AEC, LLM, ChatGPT, Gemini, compliance to building regulations, decision-making, stakeholders, AI, data expansion.*

REFERENCE: *Al Nama, E., Mahmud, M., & Al Madhoob, H. (2026). Enhancing regulatory compliance in AEC industry via LLM-powered decision frameworks. Journal of Information Technology in Construction (ITcon), 31, 201-224. https://doi.org/10.36680/j.itcon.2026.009*

1. INTRODUCTION

The AEC industry is often characterized by its complexity and fragmentation, involving multiple stakeholders, including architects, engineers, contractors, and clients. This fragmentation can lead to significant communication barriers and coordination issues, which are major sources of inefficiency (Z. Wang et al., 2020). According to (Azzouz & Papadonikolaki, 2020) AEC sector is slow to embrace and implement technological innovations. These inefficiencies are further aggravated by the industry's reliance on manual processes and the lack of integration of digital technologies (Z. Wang et al., 2020).

The AEC industry faces challenges such as ethical concerns, safety issues, and environmental impacts. Misrepresentations and lack of compliance with regulations can exacerbate these issues (Moodley et al., 2008). Non-compliance with regulations can lead to legal disputes, project delays, and increased costs. It can also damage reputations and lead to safety hazards (Moodley et al., 2008; Rubin, 2010). By leveraging data analytics, stakeholders can gain insights into complex data sets, uncover hidden patterns, and identify trends and outliers. Artificial Intelligence (AI) and big data technologies offer opportunities for improving efficiency, safety, and security in construction projects. These technologies can be applied to activity monitoring, risk management, and resource optimization, although challenges remain in their integration (Munawar et al., 2022). Automation in construction will require big data, deep learning, and ML tools. Additionally, big data, Building Information Modeling (BIM), and cloud-powered simulations can minimize project waste and improve construction quality (Munawar et al., 2022). One of the significant challenges in the AEC industry is the integration of advanced technologies. The industry's digital transformation is often slow and fragmented, with many firms struggling to implement digital tools effectively (Bhattacharya & Momaya, 2021). While the integration of these technologies holds great potential, challenges such as data fragmentation, noise, and occlusions must be resolved. A unified platform is essential for effective construction management (Z. Wang et al., 2020).

1.1 Research problem

Stakeholders in the AEC industry often struggle to navigate the complex web of building regulations necessary for regulatory compliance. These regulations are detailed and require precise interpretation to ensure that construction projects adhere to the required standards. A key challenge is the unavailability of critical information now when design decisions need to be made. Architects and engineers frequently face delays because they must consult with experts who specialize in specific chapters of the building regulations. Ensuring compliance with these intricate regulations is crucial for the safety, sustainability, and legality of construction projects, yet the process is fraught with difficulties that impede timely and informed decision-making.

From January to September 2024, the Municipalities in the Kingdom of Bahrain issued 12,205 permits for projects ranging from new builds and extensions to demolitions, fencing, and excavation. At the same time, 2,375 violations were identified which equal to 19.5% of the issued permits (The Daily Tribune, 2024).

1.2 Reserch questions

To address the identified research problem, this study formulates specific research questions that guide the investigation. These questions are designed to explore the development and effectiveness of the AI-based Q&A model in optimizing decision-making processes within the AEC industry, particularly concerning regulatory compliance under the constraints of limited data and expert availability.

- **Primary research question**
    1) How can an AI-based Q&A model be developed to optimize decision- making processes for AEC stakeholders by providing instant, accurate answers to regulatory questions during the design stage?
- **Secondary research questions**
    1) What are the key factors contributing to the selection of a base AI model to create a customized Q&A model?
    2) What methods can be employed to develop a robust AI model that operates effectively with limited data and integrates expert knowledge?

3) How can the tacit knowledge of Bahraini building code experts be codified into a structured, accessible system for use by stakeholders?

4) How effective is the AI-based Q&A model in ensuring compliance with Chapter (14) related to regulating the practice of engineering professions of the Bahraini building code?

These research questions aim to comprehensively address the challenges of regulatory compliance in the AEC industry and evaluate the potential of AI technologies, particularly the AI-based Q&A model, in overcoming these challenges. By systematically investigating these questions, the study seeks to provide actionable insights and practical solutions that can enhance decision- making processes and improve project outcomes.

1.3 Research objectives

This research seeks to address the challenges of ensuring regulatory compliance within the AEC industry in the Kingdom of Bahrain, focusing on the complexities of building regulations, limited data availability form ML training, and scarce expert knowledge. The key objectives are:

- To develop a data-driven Q&A model for building code compliance: this objective involves designing, evaluating and testing a sophisticated AI based Q&A model that assists AEC stakeholders by providing instant, accurate answers to regulatory questions.
- To evaluate the feasibility of AI models with limited data: this objective involves assessing the viability of developing an effective Q&A model using a limited dataset, augmented by human-AI collaboration.
- To codify expert knowledge into a structured system: this objective requires a systematic capture and structure of the tacit knowledge of building code experts, making it accessible for broader utilization by stakeholders.
- To develop a comprehensive framework for AI-based regulatory compliance models: this objective is to formulate a robust framework for the development of custom Q&A models for regulatory compliance, ensuring adaptability to various chapters of the building code and potential integration with BIM systems.

1.4 Research significance and contributions

This research significantly advances the AEC industry by addressing key challenges in regulatory compliance and decision-making inefficiencies. By pioneering the application of AI-based Q&A models to building code compliance, particularly in the Kingdom of Bahrain, this study introduces a novel solution to reduce costly delays and legal issues associated with non-compliance. The methodology demonstrates the feasibility of creating effective AI models even in data-scarce environments, utilizing human-AI collaboration to generate additional data and enhance model accuracy. A critical contribution lies in the systematic conversion of tacit expert knowledge into explicit, actionable insights, ensuring accessibility for AEC stakeholders. Furthermore, the proposed methodological framework supports seamless integration with BIM systems, streamlining compliance processes while optimizing project outcomes. Finally, by leveraging interdisciplinary theories, this research lays a strong foundation for future advancements in regulatory compliance and decision-making systems within the industry.

1.5 Scope of the study

This research is set within the context of the AEC industry in the Kingdom of Bahrain where the Building Permit Code - revision v1.3 (Building Code) serves as a comprehensive guide for building permit regulations within the Kingdom of Bahrain, comprising 17 chapters that address various aspects of building construction, safety, and compliance. Each chapter outlines specific requirements enforced by 17 different authorities and directorates. Notably, Chapter 14 focuses on the regulation of engineering professions (Benayat, 2025).This study will concentrate on Chapter 14, with the aim of developing an effective framework for compliance. Once this approach is validated, it has the potential to be extended and applied to other chapters within the code, addressing broader regulatory challenges.

1.6 Paper organization

This paper consists of five sections, beginning with an introduction to the study, followed by a literature review that examines relevant existing research. The methodology chapter outlines research design, data collection, and analysis methods. The findings are presented and discussed in relation to the study's objectives, and the final chapter summarizes key insights and provides recommendations for future research and practical applications.

2. REVIEW OF LITERATURE

Building regulation compliance in the AEC industry is becoming increasingly complex, requiring careful decision-making and collaboration among stakeholders. Recent research highlights the importance of Decision Support Systems (DSS), AI-driven tools, and automation in improving compliance and streamlining regulatory processes.

This section explores key advancements in stakeholder decision-making, compliance strategies, and the role of emerging technologies like Q&A models and large language models (LLMs) in enhancing regulatory adherence.

2.1 Optimization of AEC stakeholder decision-making for building regulation compliance

Decision Support Systems (DSS), defined as computer-based systems that assist in decision-making by integrating data, analytical tools, and models, are particularly valuable in this context. These systems provide structured support for analyzing complex problems and exploring alternative solutions, making them highly relevant to industries like AEC, where project success depends on balancing multiple variables and stakeholder needs (Power, 2002).

Moreover, understanding the critical factors that influence stakeholder engagement in design optimization—such as the role of subcontractors and the mechanisms for performance evaluation—is essential for achieving management goals in Engineering, Procurement, and Construction (EPC) projects (Chen et al., 2023; Muthumanickam et al., 2023). A methodology that integrates stakeholder preferences early in the design process can lead to outcomes that better reflect the needs and goals of all parties involved (Zhilyaev et al., 2022). Collectively, these approaches highlight the need for integrating stakeholder perspectives and managing uncertainties to improve decision-making in AEC projects (Li et al., 2022).

To enhance compliance with building regulations, integrating advanced DSS and performance-based metrics is crucial. (Cano et al., 2017) propose a DSS framework that supports stakeholder dialogue and dynamic risk assessment, essential for managing building infrastructure complexities and ensuring regulatory compliance. The model uses a two-stage, dynamic stochastic optimization approach with moving random time horizons, allowing for the modeling of extreme events and structural changes from stakeholder dialogue. Similarly, (Mateus et al., 2024) emphasize structuring regulations through decision analysis practices and advocate for performance-based metrics that can adapt to regulatory changes. Their work, which was validated through a case study on light rights regulation, highlights the need for clear compliance thresholds and robust monitoring mechanisms. Additionally, (Strobbe et al., (2012) discuss the benefits of integrating building simulation tools with heuristic design methods to improve compliance with energy performance regulations. This blend of stakeholder engagement, structured decision-making, and simulation tools provides a comprehensive strategy for optimizing compliance in building design and regulation processes (Strobbe et al., 2012).

However, while DSS frameworks have been widely applied in the AEC sector, gaps remain in their integration with real-time, AI-driven Q&A models, particularly in regulatory compliance contexts. Such integration could enhance the ability to address dynamic stakeholder inquiries and adapt to regulatory changes more effectively, ensuring a more proactive approach to compliance management.

2.2 Compliance with building regulations strategies

Recent studies highlight that regular and targeted training significantly improves compliance rates and reduces the likelihood of violations on construction sites (AMI Environmental, 2025).

The engagement of specialized compliance consultants has become a widespread practice. According to (K2 Integrity, 2023), the use of independent third-party consultants not only enhances compliance but also reduces legal risks and project delays (K2 Integrity, 2023).

The integration of advanced monitoring technologies is another key strategy. These technologies include the use of real-time data collection tools, automated inspection systems, and predictive analytics to monitor compliance in real-time. The implementation of such technologies is increasingly recognized as a best practice in the industry, as it allows for more proactive management of compliance-related risks (Compliance Chain, 2024).

Recent legislative changes, such as the introduction of Carlos's Law in New York, have significantly raised the stakes for non-compliance, with stricter penalties and enhanced enforcement measures. These changes aim to compel companies to prioritize compliance and adopt more robust safety protocols (K2 Integrity, 2023).

The integration of AI-driven Q&A models within BIM frameworks could offer a real-time decision support system that enhances stakeholders' ability to make informed decisions during the early design stages (Fadoul et al., 2020).While current strategies are effective, more research is needed to assess their long-term impact on reducing violations and improving overall building safety. The use of Q&A models in BIM has shown promise, but more empirical studies are needed to validate their effectiveness over time (Villaschi et al., 2022).

2.3 Integration of advanced technologies and decision support systems

While recent advancements in automated compliance checking (ACC) for building regulations have shown potential for improving both the efficiency and accuracy of compliance processes, there are still significant challenges to address. (Z. Zhang et al., 2022) conducted a systematic review of ACC technologies and identified key areas needing further research, such as improving the representation of complex regulatory rules and developing real-time DSS for stakeholders. Their findings suggest that future research should focus on creating methodologies that can effectively interpret ambiguous rules and automate the compliance process more comprehensively (Z. Zhang et al., 2022).

Moreover, there is a growing need for DSS that can provide immediate, accurate responses to inquiries from stakeholders such as architects and engineers. (Zhong et al., 2020) propose a question-answering system based on deep learning that is designed to retrieve regulatory information quickly and accurately, which is critical for rapid decision-making in compliance scenarios. By combining Natural Language Processing (NLP) with deep learning, this system could significantly enhance how regulatory compliance is managed in the AEC industry, offering precise, real-time answers to complex regulatory queries (Zhong et al., 2020).

These studies underscore the importance of continuing to explore the integration of advanced technologies and DSS within the AEC industry. Addressing these research gaps could significantly enhance stakeholder decision-making and compliance with building regulations, making these processes more efficient and timelier.

2.4 The rise of Q&A models and improved information access in AEC

Early research laid the groundwork for contemporary Q&A models in AEC. Studies explored utilizing information retrieval systems (Nabavi et al., 2023) to navigate vast amounts of project data stored in BIM software. Other research investigated the potential of natural language interfaces for BIM (N. Wang et al., 2022), paving the way for Q&A interaction with project models. These initial efforts established the importance of facilitating user-friendly access to project information.

Recent advancements in AI and NLP have led to the development of robust Q&A models specifically designed for the AEC industry. These models can access and process information from various sources, including BIM models, project documents, and industry regulations (J. Kim et al., 2022). This allows stakeholders, regardless of their technical expertise, to ask questions in natural language and receive relevant and timely answers (N. Wang et al., 2021). This democratization of access to information empowers stakeholders to participate more actively in decision-making processes (Kovacevic et al., 2008).

2.5 Enhancing regulatory compliance and informed decision-making through Q&A models

Q&A models can significantly enhance stakeholder decision-making by providing real-time insights throughout the project lifecycle. For instance, architects can use Q&A models to explore the cost implications of design choices before finalizing plans (Yuxia & Ruonan, 2021). Similarly, engineers can leverage the model to assess the structural feasibility of design variations (Robert et al., 2006). Stakeholders can also gain insights into potential environmental impacts or code compliance issues associated with design decisions (Haitao et al., 2014). These models can foster informed decision-making and reduce the risk of costly errors later in the project.

2.6 Focus on recent research related to functionalities of Q&A models for AEC applications

Recent research delves deeper into the specific functionalities of Q&A models for AEC applications. Studies explore the use of deep learning techniques to train Q&A models on domain-specific data, improving their accuracy and ability to understand complex technical inquiries. For instance, (Wu et al., 2022) proposes a novel approach for BIM object classification using machine learning algorithms. This method achieves an accuracy of 99.6% F1-measure, significantly improving the ability of Q&A models to answer intricate questions about building components. Additionally, (Xu et al., 2018) investigates the use of rule-based systems and machine learning for automating complex tasks in BIM. Furthermore, research investigates a method and system for automatic question answering generates answer data from product and service information. (Wenwu, 2018). This integration is crucial for user adoption and ensuring that stakeholders leverage Q&A models throughout the project lifecycle. By combining question-answering capabilities with the ability to generate new design options based on user queries, Q&A models could become even more powerful tools for creative exploration and decision-making in the AEC industry (Maureira et al., 2021).

2.7 Advancements and applications of prominent Large Language Models

In recent years, large language models (LLMs) have witnessed significant advancements, with organizations such as OpenAI and Google developing unique models tailored for diverse applications. Each LLM type brings distinct architectural features and capabilities, addressing various needs from general natural language processing tasks to more specialized, domain-specific applications.

- **OpenAI's GPT series (Generative Pre-trained Transformer)**

  OpenAI's GPT series, most notably GPT-4 and GPT-4o, represent some of the most influential LLMs in the field. These models excel in generating coherent, human-like text and have been widely adopted across industries. Recently, OpenAI introduced Custom GPTs, enabling users to customise models for specific tasks by integrating proprietary datasets and APIs. This allows for high precision in specialized domains such as legal and medical industries. OpenAI has also launched the GPT Store, providing a platform for users to share and monetize their custom GPT creations (OpenAI, 2024).

- **Google's Gemini**

  The Gemini series is Google's latest innovation in large language models, succeeding earlier models such as BERT, T5, and PaLM. Unlike its predecessors, Gemini is a multimodal model, capable of processing not only text but also images, audio, and video, which broadens its applicability across various industries. These models are employed in systems like Bard and Duet AI, enhancing capabilities in reasoning, complex problem-solving, and long-context comprehension Google Gemini (Google, 2024b). In addition, Gemma, a more lightweight and accessible version of Gemini, is designed for developers seeking customizable, smaller-scale models (Google, 2024c).

The features of the above LLMs can support the development of specialized Q&A models for AEC industry stakeholders. By adapting LLMs with domain-specific data and incorporating iterative reasoning techniques, these models can provide instant and accurate answers to complex engineering queries. (Shao et al., 2023) focus on enhancing AI models through continuous feedback loops, leveraging LLM evaluation metrics. LLMs are versatile tools that can be customized to excel in specific areas and suit diverse objectives (L. Zhang et al., 2024). Recent advancements, such as OpenAI's customise service, allow users to customize LLM behavior by training them on new datasets. This process adapts the model's weight to the specific domain and task, making it more focused, accurate, consistent, or creative. Users can configure their models using OpenAI's Python library or web interface (API platform & OpenAI, 2023).

2.8 Related Research

The application of LLMs to building code compliance has emerged as a promising research direction in recent years. J. Zhang (2023) offered one of the earliest explorations by examining ChatGPT's ability to convert regulatory requirements into executable computer code. In comparing its performance with advanced semantic rule-based approaches, J. Zhang identified clear limitations in accuracy but emphasized the potential of LLMs to accelerate the implementation and scaling of automated compliance systems (J. Zhang, 2023).

Expanding on this foundation, Nakhaee, et al. (2024) introduced a hybrid framework that integrates knowledge graphs with LLM reasoning. Their approach unifies structured regulatory data and natural language interpretation within a hybrid knowledge graph structure, enabling more accurate and scalable compliance checking. A case study demonstrated the feasibility of this method, highlighting both its promise and areas requiring refinement (Nakhaee et al., 2024).

In parallel, Fuchs, Witbrock, Dimyadi, and Amor (2024) investigated the translation of building regulations into machine-readable formats such as LegalRuleML. Using few-shot learning with GPT-3.5, they showed that even minimal examples could guide the model to capture the structural requirements of LegalRuleML. Their study further explored strategies such as chain-of-thought reasoning and self-consistency, demonstrating how contextualization can elicit embedded domain knowledge to support automated compliance checking (Fuchs et al., 2024).

Most recently, Madireddy et al. (2025) advanced this line of inquiry by embedding multiple LLMs, including ChatGPT, Claude, Gemini, and Llama, within Revit BIM workflows. Their system interprets building codes, generates Python scripts, and performs semi-automated compliance checks. Case studies on residential and office projects revealed significant efficiency gains, with the system reducing manual effort, improving accuracy, and streamlining the identification of regulatory violations. This work illustrates how multi-LLM integration within BIM environments can simplify complex regulatory processes and enhance reliability in compliance verification (Madireddy et al., 2025).

3. METHODOLOGICAL FRAMEWORK

The research introduces a comprehensive methodology for developing and validating a custom GPT model to optimize decision-making for stakeholders in the AEC industry. The approach integrates quantitative and qualitative methods, systematically evaluating LLMs using a human-AI generated dataset based on the Law Regulating the Practice of engineering professions in the Kingdom of Bahrain (Engineering Law). Key performance metrics such as accuracy, precision, recall, and F1 score ensured the model's technical robustness and contextual relevance. Validation through expert feedback and real-world testing provided critical insights, aligning the model's outputs with regulatory requirements and assessing its accuracy, efficiency, and user satisfaction. By combining descriptive statistics and thematic analysis, the research established a solid foundation for evaluating, refining, and validating a practical, data-driven tool tailored to industry needs. The research methodology for developing the Q&A model follows a structured 10-step process across 3 phases, ensuring a systematic and iterative approach.

Phase 1 focuses on adapting pre-trained LLMs, starting with data compilation to create a Q&A dataset on engineering law using public documents and expert insights. Next, leading LLMs (ChatGPT-4 and Gemini1.5) were evaluated using metrics like accuracy, precision, recall, and F1 score to establish baseline performance. A selection framework combining technical metrics and practical considerations identified the most suitable model, which were then subjected to fintuning (Gemini 1.5) and customisation (ChatGPT-4) to improve its domain-specific accuracy. In Phase 2, evaluated and refined thses adapted models. Performance was validated using consistent metrics, and the best-performing model (CustomGPT) was iteratively enhanced with parameter adjustments and additional data. Domain experts tested the refined model, providing specialized feedback to address gaps and improve relevance by integrating tacit knowledge. Phase 3 tested the model with real-world users, who assessed its usability and effectiveness by posing domain-specific questions. The impact on decision-making was analyzed by measuring improvements in speed, confidence, accuracy, and user outcomes. This methodology, integrating quantitative metrics and qualitative insights, ensured the development of a robust, high-performing GPT model tailored for engineering law with continuous refinements for practical usability and accuracy.

Figure (1) outlines the methodological steps applied in the research.

3.1 Dataset

The dataset used in this study is based on chapter 14 of the Bahraini Building Permit Code (v1.3), the Law that regulates the practice of engineering professions in the Kingdom of Bahrain forms the base of Chapter 14 (Benayat, 2025; CRPEP, 2014), the law comprising 45 articles across 11 clauses, the dataset outlines regulatory requirements for engineering practices. The initial phase involved extracting text from the document, followed by cleaning and standardization using NLP techniques, including tokenization, stemming/lemmatization, and stop word removal. The analysis adhered to the 5Vs of big data—Volume, Variety, Velocity, Veracity, and Value—despite its small size, it combines structured and unstructured content across and requiring advanced NLP techniques for processing. While static, it remains reliable and accurate but may need updates to address ambiguities. Its high value lies in providing rich, relevant information essential for generating Q&A pairs, making it suitable for big data applications due to its variety, veracity, and potential for augmentation ensuring the dataset's relevance and accuracy for generating a Q&A model (Abdelkader & Ibrahim, 2023; AR Paiva & T Tasiden, 2013).

Due to the dataset's small size and lack of readily available FAQs, a collaborative AI-human approach was used to expand it. ChatGPT-4 and Gemini generated diverse Q&A pairs from each article, including factual, open-ended, hypothetical, theoretical, comparative, and critical-thinking questions. Human experts reviewed and refined these AI-generated outputs to ensure thematic relevance and factual accuracy. The final dataset consisted of 235 questions, balancing human expertise with AI capabilities to create a robust Q&A model. The approach applied in this step was elaborated by Al Nama et al. (2024), who leveraged human expertise and AI for engineering regulatory data expansion in their case study with ChatGPT (Al Nama et al., 2024).

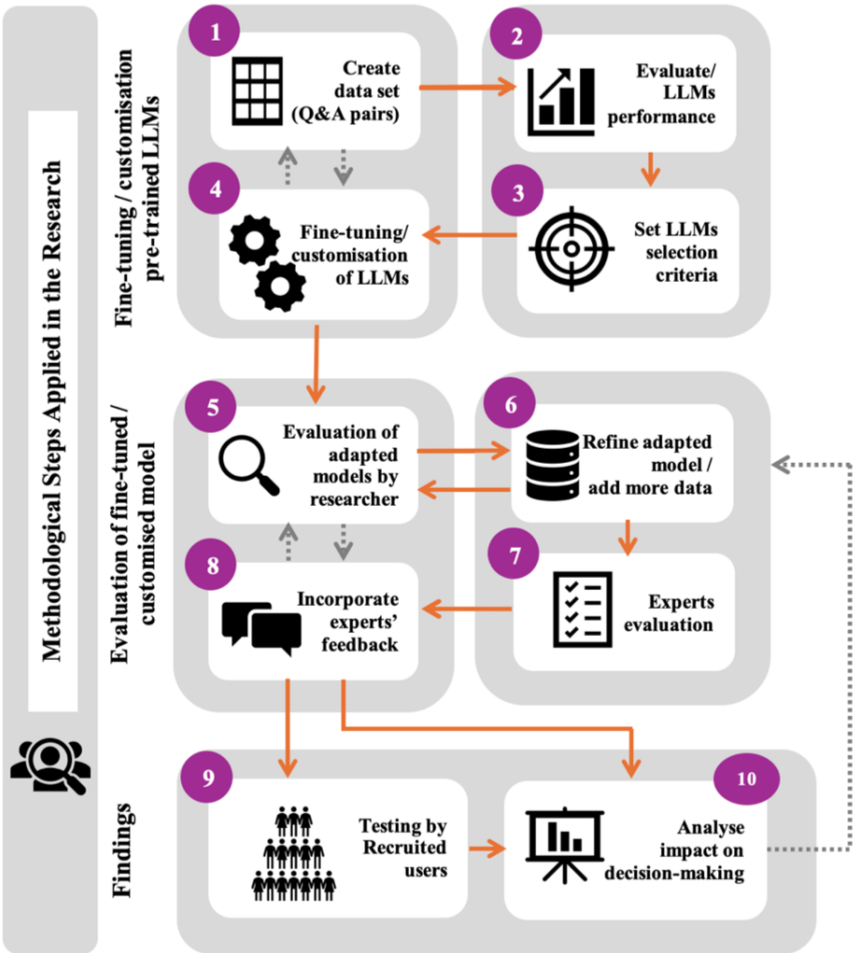

Figure 1: Methodological Steps Applied in the Research.

3.2 Step (2): Evaluate the performance of LLMs

This step involved a systematic evaluation of ChatGPT-4 and Gemini 1.5 to assess their capabilities and limitations using a combined qualitative and quantitative approach. Both models were tested against the dataset questions and meticulously graded on a scale: ‘0’ for incorrect responses, ‘0.5’ for partially correct responses, and ‘1’ for fully correct answers. The evaluation also noted whether the models supplemented their answers with additional information or referenced previous responses, providing insights into their contextual awareness and knowledge integration capabilities. The approach applied in this step was elaborated by Al Nama & Mahmud (2024), who analyzed the performance of ChatGPT-4 and Gemini by Questions Category in Engineering Regulations (Al Nama & Mahmud, 2024).

To quantify performance, key metrics—accuracy, precision, recall, and F1 score—were applied. Accuracy measured the overall proportion of correct answers, indicating model performance (D. Manning et al., 2009). Precision evaluated the proportion of true positive answers, which is crucial in legal contexts to avoid misunderstandings, while recall assessed the model's ability to retrieve all relevant answers, ensuring completeness in queries (Jordan et al., 2006). The F1 score, as the harmonic mean of precision and recall, provided a balanced metric for evaluating both accuracy and completeness (Scikit-Learn, 2024).

3.3 Step (3): Establishing selection criteria for optimal LLMs in adaption

This step established a robust framework for selecting the most suitable Large Language Model (LLM) to meet the regulatory requirements of engineering law in Bahrain. The objective was to empirically compare ChatGPT-4 and Gemini 1.5, evaluating their capabilities for domain-specific Q&A tasks. Initially, ChatGPT-4 and Gemini 1.5 were evaluated based on cost, API availability, user-friendliness, case studies, continuous development, and performance metrics (accuracy, precision, recall, F1 score).

Advancements in LLM technology prompted a shift to custom GPTs and Gemma. OpenAI’s custom GPTs allowed adaption using domain-specific datasets, enhancing relevance and accuracy for specialized fields like engineering law (OpenAI, 2024). Concurrently, Google’s Gemma, an open-source, lightweight model, offered computational efficiency and customization potential for regulatory Q&A systems (Google, 2024a, 2024b, 2024c).

This shift from general-purpose models to customizable solutions highlighted the importance of tailoring LLMs to meet regulatory compliance and practical usability, ensuring the selected model effectively addresses the unique requirements of engineering law.

3.4 Step (4): Adapting LLMs (fine-tuning and customisation)

Building on the selection criteria findings, this step involved finetuning Gemma and customising CustomGPT to create specialized Q&A models for regulatory inquiries in engineering law. The goal was to align the models with Bahrain's legal framework, ensuring accurate, context-specific, and reliable responses.

3.4.1 CustomGPT

The cutomisation of CustomGPT leveraged OpenAI’s advancements, allowing for domain-specific customization (OpenAI, 2023b). This involved integrating engineering law datasets, enhancing precision, and defining capabilities like referencing legal provisions, offering context, and handling specific user requirements.
Steps to Create CustomGPT: based on OpenAI (2023), the following steps were taken to refine the model

- Gathered relevant documents, including engineering law, prime minister edicts, and circulars, to build a comprehensive dataset.
- Cleaned, structured, and validated the data for accuracy to ensure effective model comprehension.
- Created a centralized repository of indexed documents for efficient referencing during training.
- Extracted and analyzed key provisions such as licensing requirements, fee structures, and disciplinary actions.
- Guided the model’s responses with specific instructions, ensuring clarity, transparency, and relevance.
- Trained CustomGPT on the structured and indexed dataset, focusing on legal language to align with

the engineering regulatory framework.

- Tested the model’s responses for accuracy, relevance, and comprehensiveness, with iterative feedback loops to refine outputs.
- Developed a user-friendly interface, allowing professionals to interact efficiently with the GPT while ensuring accurate referencing of legal sections.
- Suggested regular updates to reflect amendments, new regulations, and expert/user feedback, ensuring ongoing accuracy and relevance.

3.5 Step (5): Evaluation of adapted LLMs

The evaluation compared CustomGPT and Gemma's performance in responding to queries from Bahrain's engineering law. A total of 90 questions were generated, covering various types: factual, hypothetical, theoretical, open-ended, and comparative. Factual questions were prioritized due to their practical importance.

3.5.1 Evaluation design

Two questions were assigned to each article:

- One factual question to measure accuracy and precision.
- One non-factual question from the remaining types to evaluate versatility.

This resulted in a balanced dataset of 90 questions.

3.5.2 Evaluation procedure

Human experts assessed the responses generated by each model. Performance metrics—accuracy, precision, recall, and F1 score—were used to quantify the models' performance, providing insights into their strengths and weaknesses in handling legal document queries and diverse question types.

3.6 Step (6): Refine adapted model and add more data

This step acted as a bridge between Step 5 (researcher evaluation) and Step 7 (expert evaluation), ensuring the model was further optimized for accuracy and relevance. Building on the findings of Step 5, where CustomGPT was identified as the best-performing model, the focus shifted to refining the model through parameter adjustments and data enrichment to address a broader range of regulatory inquiries related to engineering law.

The refinement process involved:

- Parameter Adjustment: Iterative customisation of model parameters to enhance accuracy, efficiency, and relevance, ensuring optimal performance.
- Data Enrichment: Incorporating additional domain-specific resources, including implementing regulations, licensing forms, fee structures, amendments, and updated legal documents available online.
- Knowledge Integration: Seamlessly integrating new data with the existing knowledge base to improve the model's ability to provide contextually appropriate and accurate responses.

3.7 Th Step (7): Evaluation and continuous improvement of the adapted model through independent expert

This step involved conducting expert interviews to validate the model's performance, remove potential biases, and establish a continuous improvement mechanism through expert feedback. The interviews served three key purposes:

1) Validation: Experts evaluated the model's accuracy and relevance to ensure it aligned with real-world standards.
2) Bias Removal: Diverse expert perspectives mitigated biases from a singular evaluation.

3) Continuous Improvement: A feedback loop allowed for ongoing refinements based on expert input.

- **Participant Selection**

  Four experts were selected for interviews based on their extensive experience in legal, engineering, and regulatory domains. Their in-depth knowledge ensured comprehensive and credible feedback.

- **Interview Design**

  Each expert was provided with the model's responses to the 90 evaluation questions before the interview. The interviews were conducted individually, allowing experts to interact with the model in real-time to pose additional queries or clarify ambiguities.

- **Interview Questions**

  The interviews conducted as part of the study evaluated the model's performance across multiple dimensions to ensure a comprehensive assessment. First, general performance was reviewed in terms of accuracy, relevance, clarity, completeness, and consistency. Second, the model's behavior during real-time interaction was examined, focusing on user experience and responsiveness during live testing scenarios. Third, participants provided specific feedback, highlighting strengths, identifying weaknesses, and offering suggestions for improvement. Fourth, a technical evaluation was carried out to gauge the model's depth of knowledge, its ability to handle ambiguities, and its overall technical accuracy. Finally, additional comments were collected, capturing further observations and recommendations aimed at refining the model and enhancing its practical utility.

3.8 Step 8: Incorporation of experts' feedback

This step refined and enhanced the Custom GPT model by incorporating expert feedback to ensure its accuracy, reliability, and alignment with professional standards.

- **Data Collection and Analysis**

  Expert feedback was collected through Microsoft Forms and analyzed using thematic analysis to identify recurring themes such as accuracy, clarity, and technical depth (Braun & Clarke, 2006). Coding categorized insights into actionable areas for improvement.

- **Interview Questions**

  Experts were given access to the model's responses to a set of 90 questions and engaged with it directly during interviews to evaluate its performance. Their assessment covered several key areas, beginning with general impressions of the model's accuracy, relevance, and completeness. They also examined the consistency and clarity of its responses, identifying both strengths and areas that required improvement. Particular attention was paid to how the model handled ambiguities and the technical nuances of legal content, which are often challenging in regulatory contexts. Finally, the experts offered targeted recommendations aimed at enhancing the model's accuracy and reliability, contributing valuable insights for its refinement and future development.

- **Integration with Quantitative Data**

  Qualitative insights from expert interviews were integrated with quantitative evaluation results (e.g., accuracy, precision, recall, and F1 scores) from previous steps. This mixed-methods approach ensured a comprehensive and balanced evaluation (Creswell, 2009).

- **Continuous Improvement and Knowledge Management**

  A continuous improvement loop was established to refine the model iteratively based on expert recommendations. Both tacit knowledge (expert experience) and explicit knowledge (documented insights) were incorporated, aligning with knowledge management principles (Nonaka & Takeuchi, 1995).

- **Analytical Tools and Methods**

  Thematic analysis and qualitative data tools structured the evaluation process, while descriptive statistics quantified performance metrics (Patton, 2015).

By systematically integrating expert feedback, the model evolved into a dynamic tool capable of adapting to evolving requirements and maintaining high standards of accuracy and relevance. Results from this step are detailed in Chapter 4 of this paper.

3.9 Step (9): User testing and feedback

This step validated the model in real-world applications by testing its usability, performance, and impact on decision-making within the AEC industry while collecting detailed user feedback.

- **Participant Selection**

  Seven participants were selected from the Association of Engineering Offices based on their active engagement and familiarity with engineering law. Another random group of 7 participants from architects and engineers working in the AEC industry were also formed to gain more insights from users.

- **Study Design and Data Collection**

  Participants engaged with the Custom GPT model to answer factual questions related to key clauses of engineering law, including general provisions, licensing requirements, penalties, and professional obligations. To evaluate the model's effectiveness, users were asked to reflect on several aspects of their experience. They discussed the challenges typically encountered when navigating regulatory information through traditional methods and assessed how the model improved access to accurate and relevant content. Feedback highlighted enhanced decision-making and compliance efforts, thanks to the model's ability to deliver instant, precise, and well-referenced answers. Participants also evaluated the model's usefulness in interpreting engineering law, implementing regulations, and specific guidelines such as Chapter 14 of the Building Regulation. Finally, they considered the broader potential of applying this AI-driven approach across the entire Building Regulation framework to support more informed and optimized decisions throughout the AEC industry.

- **Data Analysis and Integration**

  Thematic analysis (Braun & Clarke, 2006) was used to identify patterns in user feedback, while qualitative tools managed the data for structured analysis. User insights were integrated into the continuous improvement loop, ensuring ongoing refinement. This process adhered to knowledge management principles to capture both tacit and explicit knowledge (Nonaka & Krogh, 2009).

This step demonstrated the model's practical utility in improving information retrieval, decision-making efficiency, and compliance efforts, while also highlighting areas for future enhancements.

3.10 Step 10: Analyzing the impact on decision-making

This final step analyzed the Custom GPT model's impact on decision-making within the AEC industry, synthesizing expert and user feedback with quantitative performance metrics.

- **Data Synthesis**

  Qualitative insights from experts and user feedback were combined with quantitative data. Statistical analysis identified correlations between model usage and improvements in decision-making outcomes.

- **Impact Assessment**

  The assessment measured the model's ability to deliver instant, accurate, and relevant answers, reducing decision-making time, improving compliance efforts, and enhancing information quality and accessibility.

- **Reporting**

Findings from the impact assessment were documented, highlighting improvements in decision-making processes attributed to the model's implementation.

- **Continuous Improvement**

  Insights from the impact analysis were integrated into the continuous improvement loop, ensuring ongoing model refinement. Feedback and new data further enhanced the model's relevance and utility.

- **Correlation with Other Methodology Steps**

  The evaluation of the model was closely aligned with other key steps in the research methodology, reinforcing its development and practical relevance. Expert interviews played a foundational role by ensuring the model adhered to technical and regulatory standards, while user feedback confirmed its applicability in real-world scenarios. These insights fed into a continuous improvement loop, enabling iterative refinement that allowed the model to evolve in response to user needs. Additionally, empirical data gathered from user testing highlighted the model's tangible impact on decision-making and compliance behavior, offering strong evidence of its utility even prior to the formal impact assessment.

3.11 Methodology review

The methodology in this research followed a structured approach to develop and evaluate the Custom Q&A model for regulatory inquiries in engineering law. It was designed across ten interconnected steps, incorporating quantitative evaluations and qualitative insights to ensure robustness and accuracy. Key phases included data compilation, model fine-tuning or csutomisation, expert validation, user testing, and impact assessment. The integration of expert and user feedback through a continuous improvement loop ensured the model evolved dynamically, aligning with real-world requirements. This comprehensive methodology established a solid foundation for assessing the model's capability to improve decision-making processes and support regulatory compliance within the AEC industry.

4. FINDINGS

This section presents the findings from the evaluation and optimization of the Custom GPT model, developed to enhance decision-making for AEC stakeholders. The findings are structured to align with the research methodology, detailing the model's evaluation. The chapter begins by assessing the baseline performance of pre-trained models, ChatGPT-4 and Gemini, in addressing complex engineering legal queries related to engineering law. These evaluations provided benchmarks for fine-tuning and customization process. Subsequent comparisons highlighted the challenges of working with limited data while demonstrating the Custom GPT's practical advantages for regulatory compliance.

Feedback from experts and users revealed the model's strengths in delivering accurate answers to straightforward queries and its limitations with interpretative or complex questions. These findings highlight the Custom GPT model's potential to significantly improve decision-making and regulatory compliance while emphasizing the need for further refinement and development.

4.1 Evaluation of LLMs performance: ChatGPT-4 and Gemini 1.5 (Step 2 of the methodology)

The evaluation criteria include accuracy, relevance, and the models' ability to handle complex queries related to engineering law and its implementing regulations. This step is foundational, as it identifies the baseline performance of the models before any fine-tuning or customization, providing a benchmark for subsequent improvements.

As this paper is part of ongoing research the results of these steps were explored in a paper presented at the ICCCBE 2024 Conference, titled "Building the Blueprint for AI-powered Compliance Checking: Analyzing ChatGPT-4 & Gemini by Question Category in Engineering Regulations" by E. Al Nama and M. Maqsood. The paper concluded that the comparative analysis of ChatGPT-4 and Gemini 1.5 in engineering law compliance revealed key insights into their strengths and areas for improvement. Both models perform well across various question types, but they differ in precision and scenario handling. ChatGPT-4 excels in critical thinking and comparison tasks, while Gemini provides more contextual information. Both models show consistent performance across different clause titles, highlighting their robustness in understanding diverse legal contexts. However, there are areas where each model can improve, suggesting opportunities for refining their answering strategies and overall performance.

4.2 Evaluation of adapted models by researcher Gemma-2b vs Custom GPT (Step 5 of the methodology)

Following the discussion in Section 3.3 of this paper, advancements in LLM technology prompted a shift to custom GPT and Gemma-2b. A set of 90 questions and answers, generated through human-AI were used to assess the performance of their performance. Accuracy, precision, recall, and F1 score were considered as analysis metrics to provide a comprehensive understanding of the strengths and weaknesses of the CustomGPT.

4.2.1 Gemma

The analysis began by testing the fine-tuned model with designated questions, but it failed to provide correct or complete answers. Despite theoretical advantages of using Low-Rank Adaptation (LoRA) Keras and Retrieval-Augmented Generation (RAG), practical implementation revealed several challenges:

- **Inaccuracy in Responses:** The model often gave short, generic answers that were inadequate for detailed regulatory document analysis, and sometimes generated irrelevant responses (acorn.io, 2024; Raj et al., 2024).
- **Hallucinations:** The model produced information does not present in the source documents, a known issue in large language models when the retrieval component fails to ground the generation process in relevant context (Raj et al., 2024).
- **Performance Benchmarks:** Metrics indicated the model's performance did not meet necessary standards for reliability and accuracy in an engineering legal context.

Given these findings, the decision was made to rule out continuing with fine-tuning Gemma 2b. The performance issues, particularly inaccuracy and hallucinations, combined with implementation complexities, suggested this approach was not effective for the project's requirements. Consequently, the researcher decided to create a custom GPT model, aiming to improve accuracy and relevance in responses by leveraging a different fine-tuning strategy.

4.2.2 CustomGPT

The evaluation of the Custom GPT model commenced by informing it that it would be presented with a set of 90 questions, as illustrated in Figure (2). These questions were delivered sequentially, allowing for a thorough assessment of the model's capabilities. The results demonstrated that the model generated correct answers for all the questions, achieving outstanding performance across key metrics: 100% accuracy, 100% precision, 100% recall, and a perfect F1 score of 100%. These findings confirm that the Custom GPT model is a highly reliable and robust tool for addressing engineering legal queries, consistently delivering precise and comprehensive responses across all evaluated dimensions.

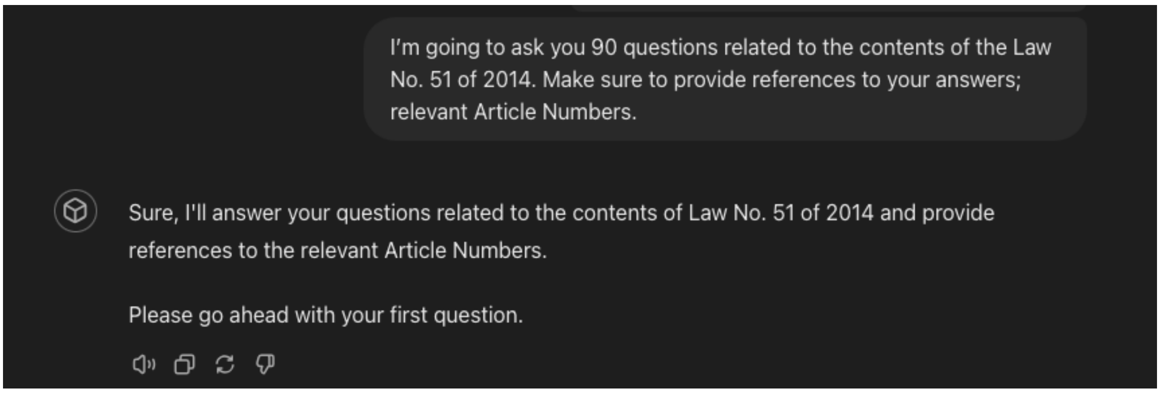

Figure 2: The Start of CustomGPT Evaluation Process.

4.2.3 Comparison of Custom GPT against base pre-trained model ChatGPT-4o performance

To assess comparative performance, the customized GPT model was evaluated against the base pre-trained ChatGPT-4o, as illustrated in Figure (4). The evaluation of ChatGPT-4o began by informing the model that it would be asked 90 questions derived from the engineering law document, shown in Figure (3). After uploading the document in PDF format, the model proceeded to answer the questions, correctly responding to 88 of them. Two questions—one factual and one comparative—were only partially answered, each earning a score of 0.5. This resulted in a total score of 89 out of 90, corresponding to an accuracy of 98.89%. The detailed performance metrics were equally strong, with precision, recall, and F1 score all recorded at 98.88%, demonstrating the model's high reliability and effectiveness in interpreting complex legal content.

4.2.4 Rationale for customization in domain-specific GPT deployment

Although the general-purpose language models GPT-4o have demonstrated remarkable performance—achieving accuracy rates as high as 98.8%—it may not be adequate in domains where precision, traceability, and control are critical. In such contexts, the use of customized GPT model becomes essential. This model can be deployed within secure infrastructures like Azure OpenAI, which offers robust safeguards for data privacy and regulatory compliance (Microsoft, 2025). Furthermore, Custom GPTs allow organizations to embed domain-specific instructions, enabling consistent control over tone, structure, and safety parameters—features particularly valuable in high-stakes environments (OpenAI, 2023a).

The importance of such control is further underscored by evidence showing that even small error margins in AI-generated outputs—such as hallucinations—can have serious implications in sensitive fields like healthcare, finance, law and engineering (Davis, 2024). Tailoring large language models to specific domains not only improves factual accuracy but also enhances contextual understanding of terminology and regulatory nuances (Si et al., 2023). Moreover, structured instruction frameworks contribute to the generation of standardized and explainable responses, which are increasingly demanded in workflows that require documentation, auditability, and legal defensibility (I et al., 2023). Collectively, these capabilities position Custom GPTs as a more reliable and controllable alternative to their general-purpose counterparts in specialized use cases.

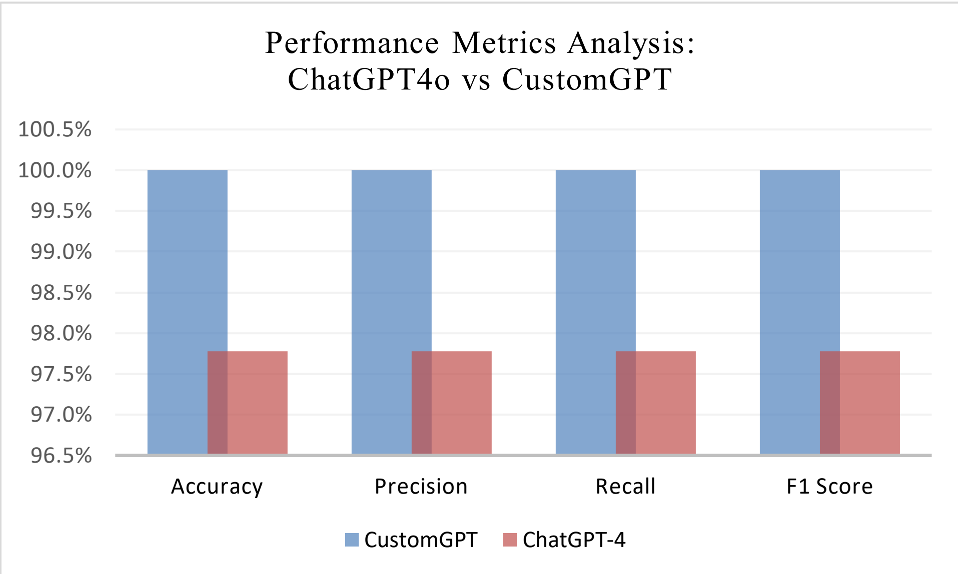

Figure 3: Comparison of Custom GPT Against Base LLM Performance.

4.3 Evaluation of the adapted model by the expert - Custom GPT (Step 7 of the methodology)

The analysis of experts' feedback revealed both the strengths and limitations of the Custom GPT model in supporting regulatory compliance in the AEC sector. Experts praised the model's high accuracy for straightforward legal queries, particularly in distinguishing between licensing requirements for engineers and architects. It was recommended for efficiently handling clear, factual questions grounded in CRPEP regulations.

However, challenges were identified in its ability to address interpretive queries and nuanced scenarios, such as evaluating complex educational qualifications or providing details on licensing upgrades. Experts noted that the model struggled with questions requiring deeper contextual understanding or involving internal, non-public decision-making criteria. This highlighted a key limitation in its reliance on publicly available data and the absence of access to dynamic regulatory updates.

The experts suggested enhancing the model by integrating explanatory notes and dynamic data sourcing from updated regulations, circulars, and executive decrees. They also recommend expanding the training data to include internal CRPEP documents to improve the model's ability to address specialized queries. Additional features, such as visual aids and voice-query functionality, were proposed to enhance usability and accessibility.

Overall, the experts emphasized the model's potential to streamline decision-making processes and regulatory compliance but stressed the importance of addressing its limitations in interpretive capabilities. They recommended a phased implementation, starting with straightforward queries, to build user confidence before handling more complex inquiries.

The feedback provided valuable insights into refining the model's performance and aligning it with the needs of stakeholders, highlighting the importance of integrating human oversight to ensure sustained accuracy and adaptability. These findings underscore the Custom GPT model's utility while emphasizing the need for continuous development to address complex regulatory challenges in the AEC industry.

The radar chart in Figure (5) shows the model's strengths in accuracy, clarity, and user interaction, but also its weaknesses with complex questions and internal information integration. The bar chart in Figure (6) highlights the need for improvements like dynamic data updates and explanatory notes for complex provisions.

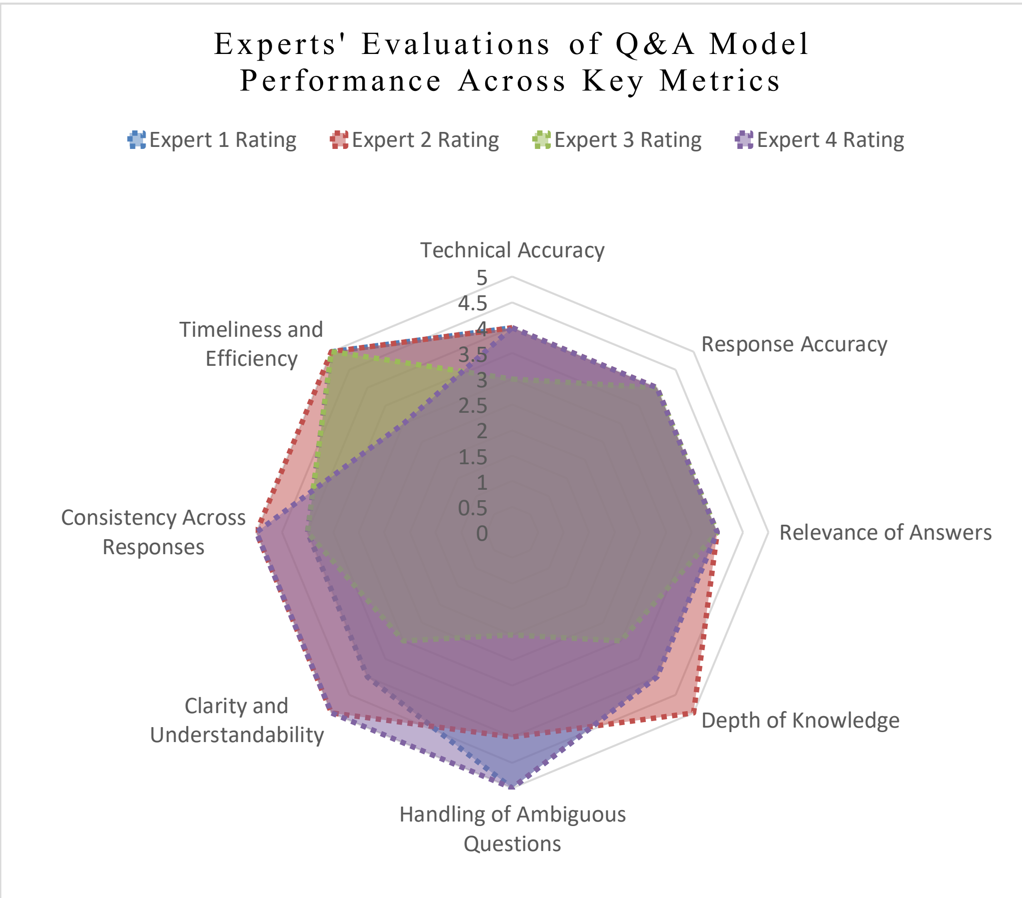

Figure 4: Experts' Evaluations of Q&A model Performance Across Key Metrics.

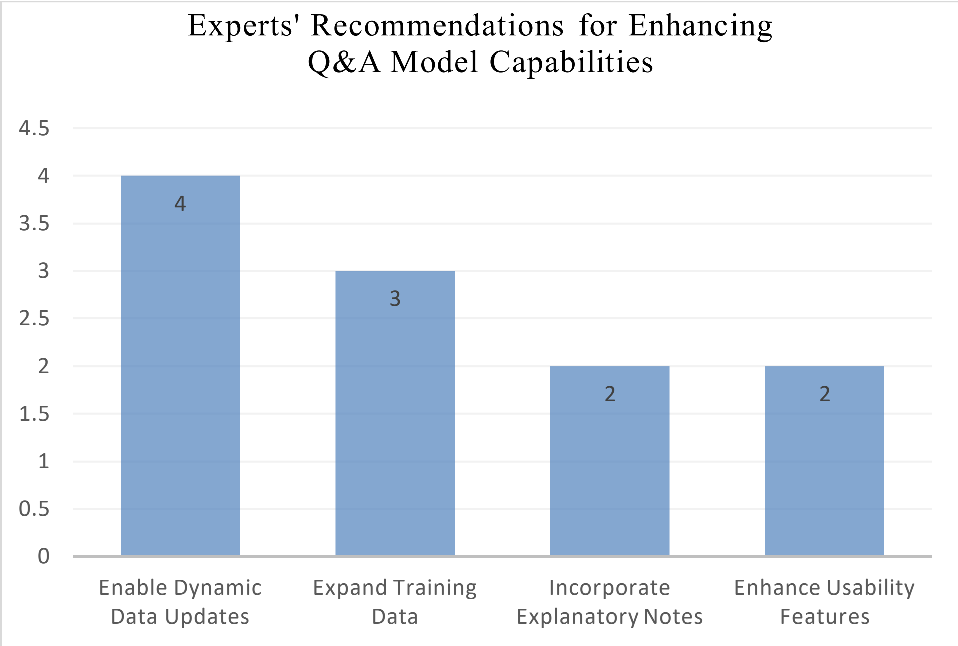

Figure 5: Experts' Recommendations for Enhancing Q&A Model Capabilities.

4.4 Analysis of the recruited users' feedback (Step 9 of the methodology)

The analysis of users' feedback highlighted the Custom GPT model's effectiveness in improving regulatory information access and decision-making in the AEC sector. Users unanimously praised their ability to deliver instant, accurate answers, significantly reducing the time and effort required to navigate regulatory information using traditional methods. Many users noted that the model's user-friendly interface streamlined their decision-making processes by providing concise, context-aware responses.

However, some users raised concerns about the model's handling of ambiguous or complex questions, particularly those requiring interpretive insights. They emphasized the importance of ensuring transparency regarding the sources of information, with several suggesting that the model include references or links for verification. Concerns were also raised about liability in cases where the model's responses led to errors in decision-making, emphasizing the need for human oversight or an option to escalate complex inquiries to a human expert.

Suggestions for enhancements included adding multilingual support (particularly Arabic), incorporating visual aids and illustrative charts, and enabling voice-query functionality to improve accessibility. Users also recommended a phased approach for deployment, beginning with straightforward queries before expanding to more complex questions as the model matures. Additionally, integrating real-time regulatory updates and ensuring consistency across responses were seen as critical to maintaining the model's relevance.

The feedback further underscored the potential of the Custom GPT model to facilitate collaboration among stakeholders by providing a centralized knowledge hub. By offering consistent regulatory references, it was noted that the model could help align discussions between clients, consultants, and regulatory bodies, reducing conflicts and improving project efficiency.

In summary, while the Custom GPT model has demonstrated strong potential to address users' needs, continuous refinement is necessary to enhance its handling of complex queries, transparency, and adaptability to the evolving regulatory landscape.

The radar chart in Figure (7) highlights feedback themes such as improved decision-making, compliance enhancement, and effective communication. It emphasizes the Custom GPT model's impact on regulatory information access and stakeholder collaboration. The bar chart in Figure (8) shows users' top recommendations: real-time updates, multilingual support, and voice integration. These insights underline the model's potential to enhance regulatory compliance and decision-making in the AEC sector. Integrating these improvements could expand the model's applicability and utility, ensuring broader adoption among stakeholders.

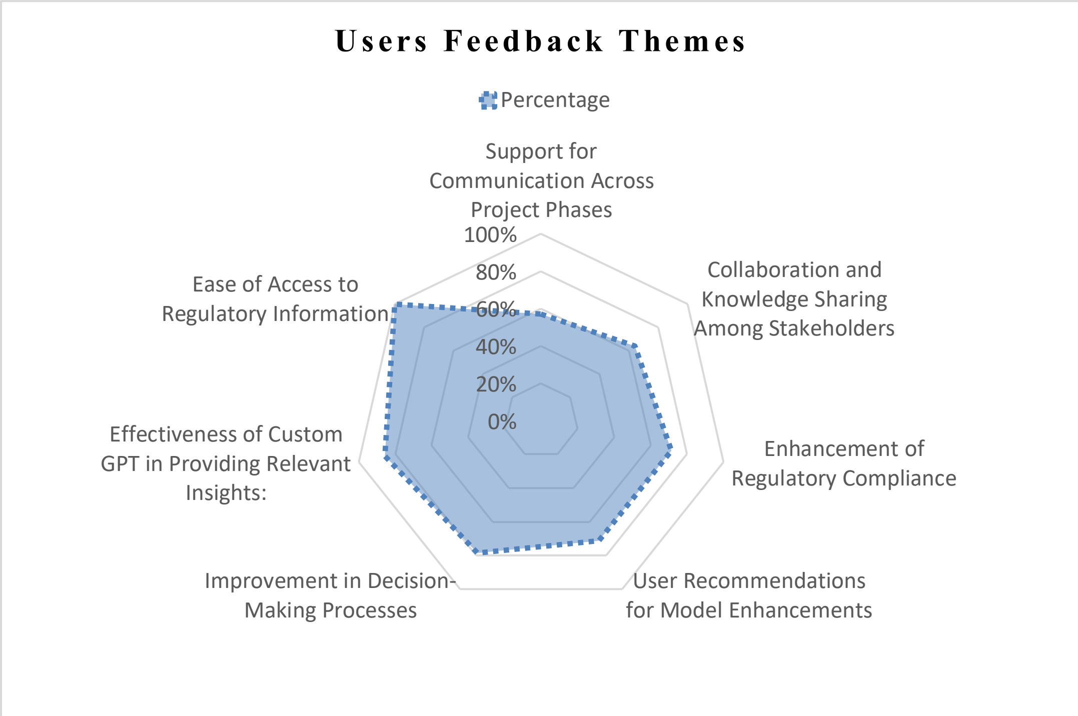

Figure 6: Users Feedback Themes.

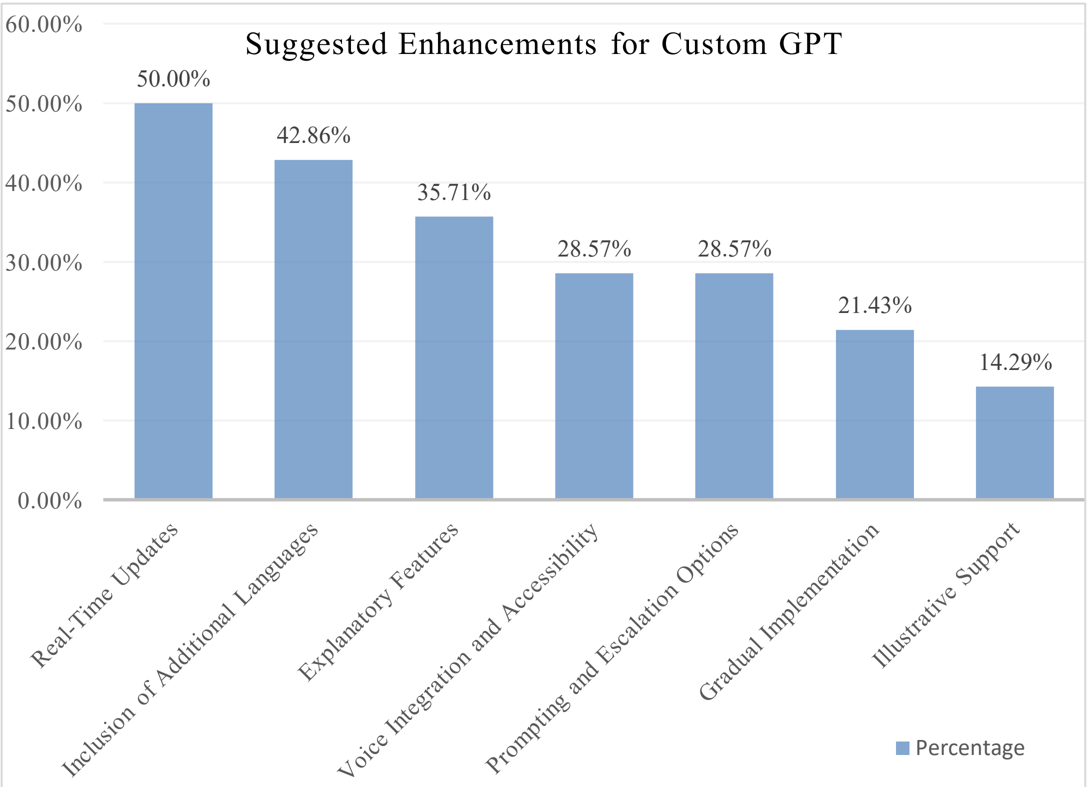

Figure 7: Suggested Enhancements for Custom GPT by Users.

4.5 Analysis of the impact on decision-making of AEC stakeholders (Step 10 of methodology)

The findings from both expert interviews and user feedback indicate that the Custom GPT model significantly improves decision-making in the AEC industry by enhancing information accessibility, decision speed, and stakeholder collaboration. Experts and users reported its ability to provide instant, accurate, and context-specific regulatory answers, reducing the need for frequent consultations and ambiguity in compliance checks. The model facilitates efficient communication among stakeholders, fosters trust through consistent responses, and dynamically adapts to regulatory updates, ensuring long-term relevance. Identified improvements include multilingual support, broader regulatory integration, enhanced source transparency, and tailored user access. These findings, summarized in Figure (9), highlight the model's impact while outlining areas for further refinement.

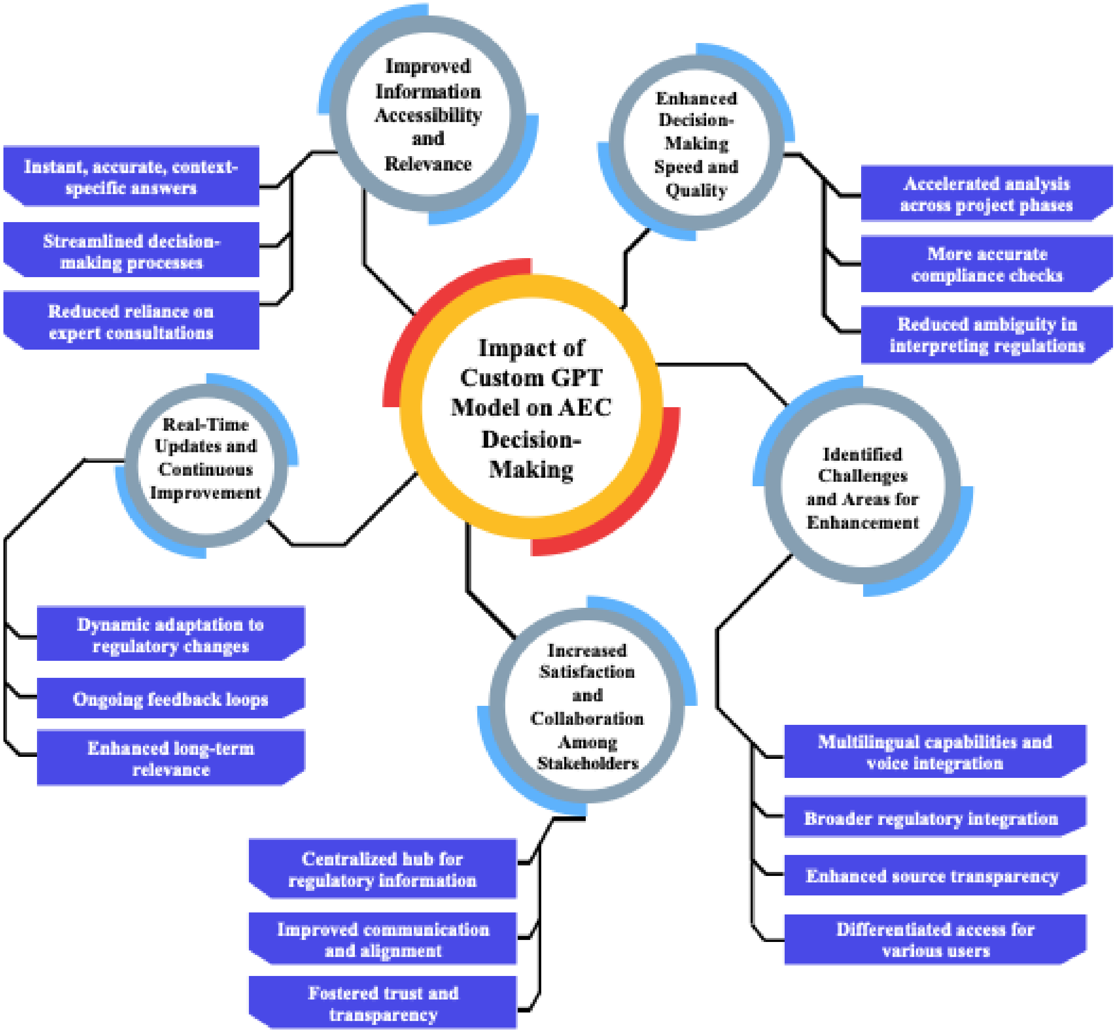

Figure 8: Key Impacts of the Custom GPT Model on AEC Stakeholders Decision-Making.

5. CONCLUSION

This study developed and evaluated a Custom GPT model to enhance decision-making and regulatory compliance in the AEC sector, particularly during the design phase. By employing a systematic 10-step methodology, the research demonstrated the model's ability to provide instant, accurate, and context-specific answers, significantly improving information accessibility, decision accuracy, and stakeholder collaboration. A notable aspect of the methodology was the integration of human-AI collaboration to expand the dataset, capturing and codifying expert tacit knowledge into explicit, structured content. This aligns with the principles of Knowledge Management

Theory, emphasizing the preservation and transformation of expert knowledge to ensure accessibility and continuity.

The framework developed is scalable and adaptable, with the potential to integrate additional building code chapters and regulatory domains. It offers a practical solution for ensuring regulatory compliance while addressing inefficiencies in traditional decision-making processes.

Guided by a systematic 10-step methodological framework, the research directly addressed the primary research question: "How can an AI-based Q&A model be developed to optimize decision-making processes for AEC stakeholders by providing instant, accurate answers to regulatory questions during the design stage?". The findings demonstrate that the model not only met the defined objectives but also provided a scalable and adaptable solution for regulatory compliance across multi-stakeholder environments. By offering accurate, instant responses to regulatory queries, the model effectively enhanced decision-making, information accessibility, and collaboration among AEC stakeholders. The study also addressed secondary research questions, shedding light on key factors for selecting a base AI model, methods for developing robust systems with limited data, and approaches to codifying expert tacit knowledge into accessible structures. The Custom GPT model's effectiveness in ensuring compliance with Chapter (14) of the building code was affirmed, while areas requiring refinement, such as handling interpretative queries and broader regulatory integration, were identified.

The study identifies a series of challenges that shape the development of the AI-based Q&A model, while also pointing toward opportunities for refinement. Questions of generalizability remain central, as regulatory frameworks differ across jurisdictions and cultural contexts, making direct transferability complex. The limited availability of published statistics on building violations further complicates validation, though professional experience and secondary sources underscore the importance of addressing this gap. Equally significant are issues of data quality, since inconsistencies or biases within regulatory datasets can influence the accuracy and reliability of the model. By acknowledging these constraints, the research clarifies the conditions under which the system can be most effective and highlights the need for adaptability in diverse environments.

Additional considerations include the technical boundaries of machine learning models, the availability of expert input, and the willingness of stakeholders to adopt new tools. Ethical concerns related to privacy and data security, alongside the rapid pace of technological change, reinforce the importance of continuous updates and responsible implementation. Resource limitations and the complexity of interdisciplinary collaboration present further hurdles, yet they also highlight the potential for innovation across fields. Taken together, these reflections position the research as a flexible and forward-looking contribution, capable of evolving with future developments while offering meaningful improvements in compliance and decision-making within the AEC industry. Future research should focus on enhancing multilingual capabilities, broadening regulatory integration, and refining tacit knowledge preservation processes. Developing adaptive AI frameworks and integrating the model with BIM systems can further enhance its utility and reliability. Additionally, addressing liability concerns and adopting phased implementation strategies will be crucial for handling complex scenarios effectively.

In summary, this research establishes a robust foundation for AI-driven regulatory compliance systems, bridging Knowledge Management Theory with advanced AI methodologies. It highlights the transformative potential of preserving and utilizing expert knowledge to optimize decision-making processes in multi-stakeholder environments, setting a precedent for future advancements in regulatory compliance frameworks.

REFERENCES

Abdelkader, B., & Ibrahim, B. (2023). Contribution to decision-making in the big data industry based on the multiparametric similarity measure for Pythagorean fuzzy sets. Journal of Logic and Computation, Volume 33, Issue 3, April 2023, Pages 517–535, 33(3), 517–535. https://academic.oup.com/logcom/article-abstract/33/3/517/6647540?redirectedFrom=fulltext

Abioye, S. O., Oyedele, L. O., Akanbi, L., Ajayi, A., Davila Delgado, J. M., Bilal, M., Akinade, O. O., & Ahmed, A. (2021). Artificial intelligence in the construction industry: A review of present status, opportunities and future challenges. In Journal of Building Engineering (Vol. 44). Elsevier Ltd. https://doi.org/10.1016/j.jobe.2021.103299 acorn.io. (2024, May 13). Fine-Tuning LLMs: Top 6 Methods, Challenges and Best Practices. Open Source Software for Building AI Assistants: Learning Center. https://www.acorn.io/resources/learning-center/fine-tuning-llm

Al Nama, E., & Mahmud, M. (2024). Building the Blueprint for AI-powered compliance Checking: Analyzing the Performance of ChatGPT-4 & Gemini by Question Category in Engineering Regulations. The International Conference on Computing in Civil and Building Engineering (ICCCBE) - 2024.

Al Nama, E., Mahmud, M., & Al Madhoob, H. (2024, December). Leveraging Human Expertise and AI for Engineering Regulatory Data Expansion: A Case Study with ChatGPT. International Conference in Emerging Technologies for Sustainability and Intelligent Systems (ICETSIS 2024). https://icetsis24.asu.edu.bh/

AMI Environmental. (2025). Compliance and Safety: Environmental Consultants' Roles in Building Projects.

API platform, & OpenAI. (2023). Transforming Work and Creativity with AI. https://openai.com/product

AR Paiva, & T Tasiden. (2013). (12) United States Patent. chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://patentimages.storage.googleapis.com/53/6a/7f/940e0ac09c2cd1/US8412651.pdf

Azzouz, A., & Papadonikolaki, E. (2020). Boundary-Spanning for Managing Digital Innovation in the AEC Sector. Architectural Engineering and Design Management. https://doi.org/10.1080/17452007.2020.1716693

Benayat. (2025). Benayat. https://www.benayat.bh/building-permits/

Bhattacharya, S., & Momaya, K. (2021). Actionable Strategy Framework for Digital Transformation in AECO Industry. Engineering, Construction and Architectural Management. https://www.emerald.com/insight/content/doi/10.1108/ECAM-07-2020-0587/full/html

Braun, V., & Clarke, V. (2006). Using Thematic Analysis in Psychology. Qualitative Research in Psychology, 3(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Cano, E. L., Moguerza, J. M., Ermolieva, T., & Yermoliev, Y. (2017). A Strategic Decision Support System Framework for Energy-Efficient Technology Investments. TOP, 2. https://doi.org/10.1007/s11750-016-0429-9

Chen, Y., Ren, Z., Hu, B., & Zheng, H. (2023). Investigation of the Critical Factors Influencing Multi-Stakeholders' Participation in Design Optimization of EPC Projects. Buildings, 13(7). https://doi.org/10.3390/buildings13071654

Compliance Chain. (2024, January 26). Compliance in Construction (Best Way in 2023). https://compliancechain.com/compliance-in-construction-best-way-in-2023/

Creswell, J. W. (2009). Research Design- Qualitative, Quantitative, and Mixed Methods Approaches. chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://www.ucg.ac.me/skladiste/blog_609332/objava_105202/fajlovi/Creswell.pdf

CRPEP. (2014). Law No. (51) of 2014 with respect to Regulating the Practice of Engineering Professions. https://www.crpep.bh/

D. Manning, C., Raghavan, P., & Schütze, H. S. (2009). Introduction to Information Retrieval.

Davis, W. (2024). Hospitals Use a Transcription Tool Powered by a Hallucination-prone OpenAI Model. https://www.theverge.com/2024/10/27/24281170/open-ai-whisper-hospitals-transcription-hallucinations-studies?utm_source=chatgpt.com

Fadoul, A., Tizani, W., & Arturo Osorio-Sandoval, C. (2020). A Knowledge-Based Model for Constructability Assessment of Buildings Design Using BIM. The 18th International Conference on Computing in Civil and Building Engineering, 147–159. https://link.springer.com/chapter/10.1007/978-3-030-51295-8_13

Fuchs, S., Witbrock, M., & Amor, R. (2024). Using Large Language Models for the Interpretation of Building Regulations. 13th Conference on Engineering, Project and Production Management. https://chat.openai.com/

Google. (2024a, April). Colaboratory (Colab). https://colab.research.google.com/

Google. (2024b, April). Gemini Official Blog. https://blog.google/products/gemini/

Google. (2024c, April). Gemma Open Models. https://ai.google.dev/gemma

Haitao, Z., Ning, G., Yong, J., Shutao, X., & Zhao, C.-Z. (2014). Automatic Question-Answering System. Journal of Information Technology. https://doi.org/10.1080/01446190801965368

I, M., Saxena, S., Prasad, S., Prakash, M. V. S., Shankar, A., V, V., Vaddina, V., & Gopalakrishnan, S. (2023). Minimizing Factual Inconsistency and Hallucination in Large Language Models. http://arxiv.org/abs/2311.13878

J. Kim, S. Chung, S. Moon, & S. Chi. (2022). Feasibility Study of a BERT-based Question Answering Chatbot for Information Retrieval from Construction Specifications. 2022 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM). https://ieeexplore.ieee.org/document/9989625/authors#authors

Jordan, M., Kleinberg, J., & Schölkopf, B. (2006). Pattern Recognition and Machine Learning. chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://www.microsoft.com/en-us/research/uploads/prod/2006/01/Bishop-Pattern-Recognition-and-Machine-Learning-2006.pdf

K2 Integrity. (2023). Construction Safety: How Robust Compliance Programs Can Ensure Worker Safety. https://www.k2integrity.com/en/knowledge/expert-insights/2023/construction-safety-how-robust-compliance-programs-can-ensure-worker-safety/

Kovacevic, M., Nie, J.-Y., & Davidson, C. (2008). Providing Answers to Questions from Automatically Collected Web Pages for Intelligent Decision Making in the Construction Sector. Journal of Computing in Civil Engineering, 22(1), 3–13. https://doi.org/10.1061/(asce)0887-3801(2008)22:1(3)

Li, X., Liu, S., & Sun, Y. (2022). A GP-Based Hierarchical Objectives Decision-Making Method for Building Energy Efficiency Optimization. Buildings, 12(1). https://doi.org/10.3390/buildings12010052

Madireddy, S., Gao, L., Din, Z. U., Kim, K., Senouci, A., Han, Z., & Zhang, Y. (2025). Large Language Model-Driven Code Compliance Checking in Building Information Modeling. Electronics (Switzerland), 14(11). https://doi.org/10.3390/electronics14112146

Mateus, R. J. G., Pinto, F. S., Fauth, J., Azenha, M., Granja, J., Veludo, R., Muniz, B., Reis, J., & Marques, P. (2024). Improving Regulations for Automated Design Checking Through Decision Analysis Good Practices: A Conceptual Application to the Construction Sector. Flexible Automation and Intelligent Manufacturing: Establishing Bridges for More Sustainable Manufacturing Systems (FAIM 2023), 160–169. https://doi.org/10.1007/978-3-031-38241-3_19

Maureira, C., Pinto, H., Yepes, V., & Garcia, J. (2021). Towards an AEC-AI Industry Optimization Algorithmic Knowledge Mapping: An Adaptive Methodology for Macroscopic Conceptual Analysis. IEEE Access, 9, 110842–110879. https://doi.org/10.1109/ACCESS.2021.3102215

Microsoft. (2025, June 24). Data, Privacy, and Security for Azure OpenAI Service. https://learn.microsoft.com/en-us/azure/ai-foundry/responsible-ai/openai/data-privacy?tabs=azure-portal

Moodley, K., Smith, N., & Preece, C. N. (2008). Stakeholder Matrix for Ethical Relationships in the Construction Industry. Construction Management and Economics , 625–632. https://www.tandfonline.com/doi/full/10.1080/01446190801965368?scroll=top&needAccess=true

Munawar, H. S., Ullah, F., Qayyum, S., & Shahzad, D. (2022). Big Data in Construction: Current Applications and Future Opportunities. Big Data and Cognitive Computing, 6(1). https://doi.org/10.3390/bdcc6010018

Muthumanickam, N. K., Brown, N., Duarte, J. P., & Simpson, T. W. (2023). Multidisciplinary Design Optimization in Architecture, Engineering, and Construction: A Detailed Review and Call for Collaboration. In Structural and Multidisciplinary Optimization (Vol. 66, Issue 11). Springer Science and Business Media Deutschland GmbH. https://doi.org/10.1007/s00158-023-03673-y

Nabavi, A., Ramaji, I., Sadeghi, N., & Anderson, A. (2023). Leveraging Natural Language Processing for Automated Information Inquiry from Building Information Models. Journal of Information Technology in Construction, 28, 266–285. https://doi.org/10.36680/J.ITCON.2023.013

Nakhaee, A., Elshani, D., & Wortmann, T. (2024). A Vision for Automated Building Code Compliance Checking by Unifying Hybrid Knowledge Graphs and Large Language Models. Design Modelling Symposium Berlin, 445–457. https://doi.org/https://doi.org/10.1007/978-3-031-68275-9_36

Nonaka, I., & Krogh, G. von. (2009). Tacit Knowledge and Knowledge Conversion: Controversy and Advancement in Organizational Knowledge Creation Theory. JSTOR, 20, 635–652.

Nonaka, I., & Takeuchi, H. (1995). The Knowledge - Creating Company.

OpenAI. (2023a). Key Guidelines for Writing Instructions for Custom GPTs. https://help.openai.com/en/articles/9358033-key-guidelines-for-writing-instructions-for-custom-gpts?utm_source=chatgpt.com

OpenAI. (2023b, November 6). Introducing GPTs. https://openai.com/index/introducing-gpts/

OpenAI. (2024, April). OpenAI. https://openai.com/

Patton, M. Quinn. (2015). Qualitative Research & Evaluation Methods (4th ed.).

Power, D. J. (2002). Decision Support Systems: Concepts and Resources for Managers. Quorum Books. chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://scholarworks.uni.edu/cgi/viewcontent.cgi?article=1066&context=facbook

Raj, M. J., VM, K., Warrier, H., & Gupta, Y. (2024). Fine Tuning LLM for Enterprise: Practical Guidelines and Recommendations. ArXiv. https://doi.org/10.48550/arXiv.2404.10779

Robert, J.-M., Moulet, L., Lizarralde, G., Davidson, C., Nie, J.-Y., & Sylva, L. Da. (2006). Finding Out: A System for Providing Rapid and Reliable Answers to Questions in the Construction Sector. Construction Innovation. https://doi.org/10.1108/14714170610713926

Rubin, R. A. (2010). Review of Legal Aspects of Architecture, Engineering, and the Construction Process. Journal of Legal Affairs and Dispute Resolution in Engineering and Construction.

Scikit-Learn. (2024). F1_score. https://scikit-learn.org/stable/modules/generated/sklearn.metrics.f1_score.html

Shao, Z., Gong, Y., Shen, Y., Huang, M., Duan, N., & Chen, W. (2023). Enhancing Retrieval-Augmented Large Language Models with Iterative Retrieval-Generation Synergy. Findings of the Association for Computational Linguistics: EMNLP 2023. https://doi.org/10.18653/v1/2023.findings-emnlp.620

Si, C., Gan, Z., Yang, Z., Wang, S., Wang, J., Boyd-Graber, J., & Wang, L. (2023). Prompting GPT-3 To Be Reliable. http://arxiv.org/abs/2210.09150

Strobbe, T., Pauwels, P., Verstraeten, R., De Meyer, R., & Campenhout, V. (2012). Optimization in Compliance Checking Using Heuristics: Flemish Energy Performance Regulations (EPR). EWork and EBusiness in Architecture, Engineering and Construction.

The Daily Tribune. (2024, November 19). Building Permit Approval Made Simple. Building permit approval made simple

Villaschi, F. S., Carvalho, J. P., & Bragança, L. (2022). BIM-Based Method for the Verification of Building Code Compliance. Applied System Innovation, 5(4). https://doi.org/10.3390/asi5040064

Wang, N., Issa, R. R. A., & Anumba, C. J. (2021). Query Answering System for Building Information Modeling Using BERT NN Algorithm and NLG. Computing in Civil Engineering 2021 - Selected Papers from the ASCE International Conference on Computing in Civil Engineering 2021, 425–432. https://doi.org/10.1061/9780784483893.053

Wang, N., Issa, R. R. A., & Anumba, C. J. (2022). NLP-Based Query-Answering System for Information Extraction from Building Information Models. Journal of Computing in Civil Engineering, 36(3). https://doi.org/https://doi.org/10.1061/(ASCE)CP.1943-5487.0001019

Wang, Z., He, B., Yang, Y., Shen, C., & Peña-Mora, F. (2020). Building a next-generation AI Platform for AEC: A Review and Research Challenges. 37th CIB W78 Information Technology for Construction Conference, 27–45. https://doi.org/10.46421/2706

Wenwu, Q. (2018). Automatic Questioning and Answering Method and System. https://typeset.io/papers/automatic-questioning-and-answering-method-and-system-2ym6e8hcjf

Wu, J., Asce, S. M., Akanbi, T., Zhang, J., & Asce, A. M. (2022). Constructing Invariant Signatures for AEC Objects to Support BIM-based Analysis: Automation through Object Classification. Computing in Civil Engineering. https://www.researchgate.net/publication/361671307_Constructing_Invariant_Signatures_for_AEC_Objects_to_Support_BIM-Based_Analysis_Automation_through_Object_Classification

Xu, S., Li, W., Tang, L. C., Yang, L., & Tang, Q. (2018). Artificial Intelligence Assisted Professional Work in BIM: A Machine Reasoning Extension. Creative Construction Conference (2108), 16–23. https://doi.org/10.3311/ccc2018-003

Yuxia, L., & Ruonan, R. (2021). Automatic Question-Answering System Based on Semi-Supervised Learning and Text-to-SQL Model. Journal of Artificial Intelligence Research. https://doi.org/10.1613/jair.1.12345

Zhang, J. (2023). How Can ChatGPT Help in Automated Building Code Compliance Checking? Proceedings of the International Symposium on Automation and Robotics in Construction, 63–70. https://doi.org/10.22260/ISARC2023/0011

Zhang, L., Jijo, K., Setty, S., Chung, E., Javid, F., Vidra, N., & Clifford, T. (2024). Enhancing Large Language Model Performance to Answer Questions and Extract Information More Accurately. ArXiv. http://arxiv.org/abs/2402.01722

Zhang, Z., Ma, L., & Broyd, T. (2022). Towards Fully-Automated Code Compliance Checking Of Building Regulations: Challenges For Rule Interpretation And Representation. European Conference on Computing in Construction.

Zhilyaev, D., Binnekamp, R., & Wolfert, A. M. R. (2022). Best Fit for Common Purpose: A Multi-Stakeholder Design Optimization Methodology for Construction Management. Buildings, 12(5). https://doi.org/10.3390/buildings12050527

Zhong, B., He, W., Huang, Z., Love, P., Tang, J., & Luo, H. (2020). A Building Regulation Question Answering System: A deep Learning Methodology. Advanced Engineering Informatics, 46. https://www.sciencedirect.com/science/article/abs/pii/S1474034620301658?via%3Dihub