Saparamadu, P.V.I.N., Jayasena, H.S. and Eranga, B.A.I., 2024. **Blueprint for a natural language processing powered nexus for** regulatory and legal landscape in construction. In: Sandanayake, Y.G., Waidyasekara, K.G.A.S., Ranadewa, K.A.T.O. and Chandanie, H. (eds). *Proceedings of the 12th World Construction Symposium*, 9-10 August 2024, Sri Lanka. pp. 306-317. DOI: https://doi.org/10.31705/WCS.2024.24. Available from: https://ciobwcs.com/papers/

BLUEPRINT FOR A NATURAL LANGUAGE PROCESSING POWERED NEXUS FOR REGULATORY AND LEGAL LANDSCAPE IN CONSTRUCTION

P.V.I.N. Saparamadu[1], H.S. Jayasena[2], and B.A.I. Eranga[3]

ABSTRACT

The recent exponential advancements in Natural Language Processing (NLP) are catalysing a paradigm shift in the world, directing the construction industry towards an era of smart construction. The proficiency of NLP in comprehending and assimilating vast quantities of human language data aligns aptly with the construction sector's exigency for enhanced management of its unstructured textual data. Given the frequent alterations in regulatory frameworks and the dispersed nature of project data, there arises a compelling need for a Natural Language Processing Powered Compliance Management Nexus (NLP-PCMN), which facilitates expedited access to consolidated information via mobile platforms. This study aims to develop a blueprint for implementing an NLP-PCMN in the construction industry. By conducting semi-structured interviews with 20 experts spanning the domains of construction and Artificial Intelligence (AI) alongside a focus group to outline the technological framework of the NLP-PCMN, the research underscores the need to implement such a system. The envisaged system is poised to address challenges such as navigating contract clauses, correspondence analysis and ensuring legal compliance with planning and building codes and legal provisions. The proposed NLP-PCMN presents a comprehensive solution integrating these features through large language models that work as a question-and-answering system. Key findings include the necessity of automating the regulatory and legal data in construction, stakeholder empowerment through NLP-PCMN, identifying the nodes of the NLP-PCMN and the technical blueprint to implement the NLP-PCMN.

Keywords: *Artificial Intelligence (AI); Construction Law; Natural Language Processing (NLP); Smart Construction.*

1. INTRODUCTION

As compelling evidence of the contemporary paradigm shift in AI technologies, the global AI in construction market size was valued at USD 696 million in 2023 and is projected to reach fivefold by 2032 (360iResearch, 2023). Among those technologies, Natural Language Processing (NLP) technology made a significant leap in the last five

[1] Research Student, Department of Building Economics, University of Moratuwa, Sri Lanka, ishini@concolabs.com
[2] Senior Lecturer, Department of Building Economics, University of Moratuwa, Sri Lanka, suranga@uom.lk
[3] Lecturer, Department of Building Economics, University of Moratuwa, Sri Lanka, isurue@uom.lk

years as AI exceeded human performance levels on basic reading comprehension benchmarks (Stone et al., 2021). NLP is a subfield of AI for difficult language-related problems, such as machine translation, question-answering and summarisation (Lauriola et al., 2022). The recent trend of NLP applications is a consequence of introducing Large Language Models (LLMs) such as OpenAI's ChatGPT, Google's BERT and MT5 (Zhang et al., 2022). LLMs are a subset of NLP, representing advanced models that emerged in 2018 (Cambria & White, 2014).

Although NLP applications in various fields began in the 1960s, their adaptation to the construction industry commenced in the 1990s (Khurana et al., 2023). The first application of NLP to the construction industry was developed in 1989, aiming to aid construction managers in retrieving vital information for decision-making (Khurana et al., 2023). NLP applications have been developed for its thematic areas in the construction industry. These include enhancing information and document management processes, improving compliance management, evaluating public perceptions of construction projects and optimising contract management (Hassan & Le, 2020).

Babatunde et al. (2023) revealed that proficient compliance management can be achieved through NLP by enhancing contract accuracy, efficiency and transparency. Legal and regulatory compliance in the construction industry is of paramount importance due to the complex and stringent laws governing this sector (Beach et al., 2015). Regulatory compliance ensures that construction projects adhere to established standards, codes and regulations. According to Marzouk and Enaba (2019), legal compliance includes contract management, appropriate documentation and correspondence procedures. Automating compliance processes in construction can streamline operations, reduce human error, and enhance efficiency (Beach et al., 2015). Furthermore, an automated compliance management system provides real-time monitoring and reporting capabilities (Beach et al., 2020). Integrating automation into legal and regulatory compliance processes in construction is essential for ensuring project success.

Parikh et al. (2023) documented the landmark court case that used NLP to form a judicial decision for the first time. It is a result of the evolution of NLP technology in recent years, driven by advancements in deep learning and machine learning (Khurana et al., 2023). One of the key breakthroughs has been the development of powerful language models, which have been trained on vast amounts of textual data that can capture the nuanced relationships between words and the context in which they are used (Zhang et al., 2022).

Yan et al. (2020) emphasised that the legal and regulatory data in the construction domain are available as unstructured data sources such as text documents. Therefore, efficient and intelligent extraction and interpretation of this textual data is vital for the cost-effective management of projects (Wu et al., 2022). NLP provides the solution for it through the analysis of text structures and words (Nadkarni et al., 2011). Thus, with NLP advancements continuing to unfold, it has become increasingly important for construction professionals to rethink how to leverage it to enhance their practices and processes.

It was found that numerous models have been developed for legal and contractual domains, utilising various advancements in NLP technology. For example, Beach et al. (2020) and Lee et al. (2023) have contributed models geared towards regulatory compliance automation, while Hand et al. (2021) and Lee et al. (2020) have focused on models for contract management. After testing these models, the above-mentioned literature found a significant increase in the accuracy of compliance management.

Although several NLP models have been developed in the literature, NLP adoption in the construction industry is still in its infant stage (Madan & Ashok, 2023). Wu et al. (2022) identified that the lack of awareness in the industry on utilising LLMs for their business purposes was a fundamental reason for the current state of adoption. Adding to the statement above, Madan and Ashok (2023) highlighted the limited adoption of NLP because there is not a widely recognised NLP model in commercial use, and the existing models aren't feasible for practical application in the construction sector.

Hence, it can be suggested that an ideal approach to harnessing the potential of NLP entails establishing a robust nexus comprising specialised tools tailored to be used by all stakeholders of the industry. This nexus-oriented strategy facilitates seamless integration and synergy among various NLP capabilities, avoiding the need for disparate toolsets (Mitchell & Mancoridis, 2006). A single platform housing all software tools is crucial for several reasons. It simplifies the management process by reducing the complexity and overhead associated with managing multiple systems. This approach also reduces the risk of data inconsistencies by storing and processing data within a single system (Mitchell & Mancoridis, 2006). Thus, there is a timely need for an NLP-powered nexus for effective adoption in the construction industry.

While ChatGPT has demonstrated significant capabilities in NLP, it is insufficient as a standalone solution for specialised applications such as legal research. ChatGPT's general-purpose design lacks the domain-specific knowledge (Parikh et al., 2023). Studies have shown that domain-specific NLP models significantly outperform general models in specialised fields (Jurafsky & Martin, 2021). Therefore, a tailored approach incorporating domain-specific NLP tools is essential.

Despite the growing interest in NLP tools, there remains a shortage of research on their application across various domains, including the construction industry. Jallan et al. (2019) studied the development of an NLP model to conduct a comprehensive survey of legal cases; however, they used statistical algorithms rather than LLMs. However, a study by Moon et al. (2022) used recent developments in NLP to review specifications. Furthermore, while recent research by Shaikh and Gohar (2024) has investigated the use of chatbots in contract management, it does not offer a comprehensive solution for the entire legal landscape. Therefore, this study seeks to pioneer such a concept of an accessible legal database tailored explicitly for the construction industry. Hence, it aims to develop a blueprint for implementing an NLP-powered Construction Management Network (NLP-PCMN). The paper begins with a literature review and continues by investigating the need for an NLP-PCMN. Finally, it presents an implementation blueprint detailing the architecture and key components necessary for its development.

2. LITERATURE REVIEW

2.1 NLP APPLICATIONS IN THE CONSTRUCTION INDUSTRY

Studies suggest that NLP has been used in the construction industry for following application scenarios, including filtering information, organising documents, using expert systems and automating compliance checking (Wu et al., 2022). From the literature analysis, 91 NLP models were identified for various domains, as illustrated in Figure 1. The models discussed in the literature were developed using obsolete advancements in NLP, such as rule-based techniques, probability models and neural networks. Among the main application areas, NLP for contract management accounted for 31%, making it the domain with the highest number of applications. Compliance checking application domain constituted 23% of the NLP models. Thus, the NLP models developed for the regulatory and legal landscape of the construction industry amounted to 54% of all models developed. Figure 1 shows the distribution of these models across the literature.

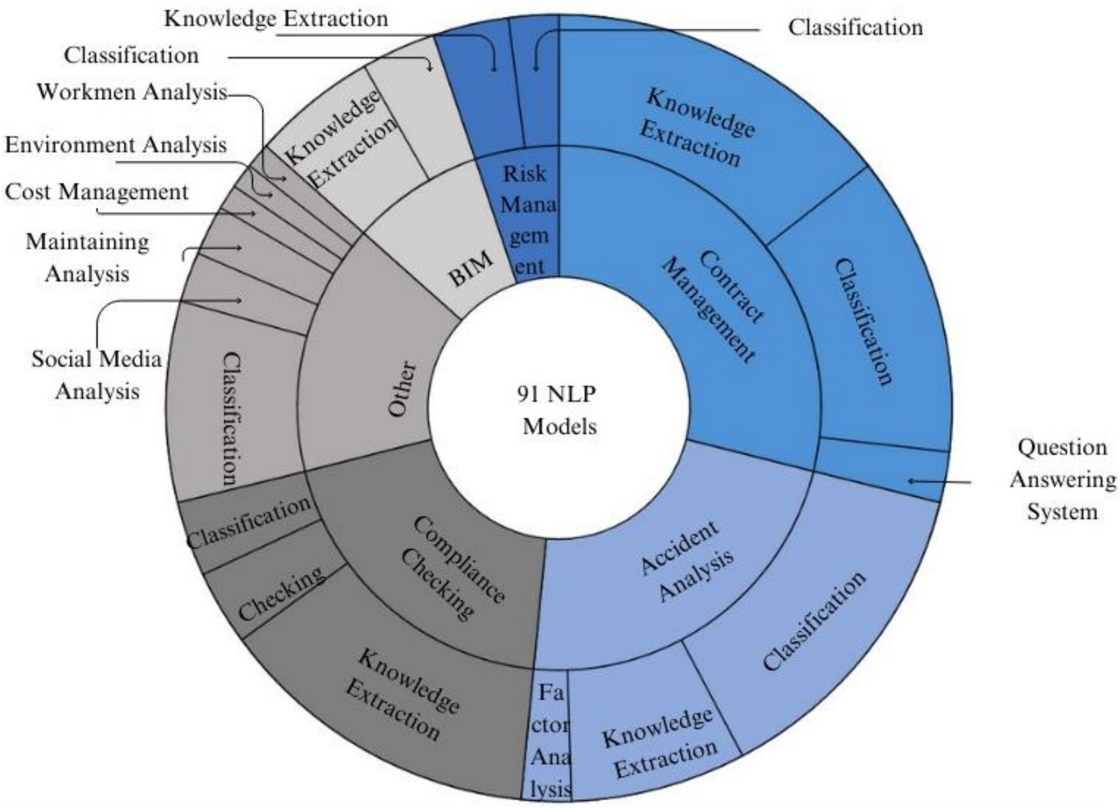

Figure 1: Distribution of NLP models across application areas (Source: Developed by authors)

2.2 STATE OF THE ART OF NLP MODELS IN CONSTRUCTION

Recent advancements in NLP have revolutionised the construction industry's approach to contract analysis and management. Notably, Padhy et al. (2021) demonstrated an 80% increase in efficiency with an NLP model designed to detect exculpatory sentences. As validated by Hand et al. (2021), F1 scores exceeding 70% indicate a reliable and effective model. Lee et al. (2019) and Lee et al. (2020) further demonstrated the competence of NLP in automatically detecting problematic clauses with impressive F1 scores of 81.8% and 80%, respectively. These findings highlight the robustness of NLP in scrutinising contractual documents, flagging critical clauses, and enhancing decision-making processes in the construction industry.

2.3 REGULATORY AND LEGAL CONSIDERATIONS IN CONSTRUCTION

To formulate a nexus facilitating efficient regulatory and legal compliance within construction projects, it is crucial to identify the textual data necessitating analysis. Through literature analysis, various types of data were identified, as outlined in Table 1.

Table 1: Sources of textual data in a construction project

Citation	Sources of Textual Data
Soibelman et al., 2008	conditions of contracts, specifications, change orders, requests for information, meeting minutes
Kelley, 2012	Statutes and ordinances, agency regulations, international treaties, case law, contract clauses
Szewc, 2022	environmental protection law, civil law regulations, public procurement regulations, property law, planning law
Murdoch & Hughes, 2002	insurance law, contract law, dispute resolution procedures, case law, standard conditions, contract data, construction service agreements, procurement law
Marzouk & Enaba, 2019	contract, variation order log, site instruction, progress reports, request for information log, cost schedule data, claim data

3. RESEARCH METHODOLOGY

Qualitative research, particularly grounded in interpretivism, stands as the optimal approach for the development of an NLP-PCMN. At the core of interpretivism lies the belief that reality is subjective, emphasising the importance of understanding the unique perspectives of the individuals involved in the data collection (Potrac et al., 2014). In the context of NLP-PCMN, this study aims to find a speculative ideal for a prospective technology that is well suited to the interpretive philosophy. By embracing qualitative methodologies, the fluidity of different experiences can be seen while shedding light on emergent patterns (Saunders et al., 2018).

The research process encompassed two phases of data collection. Initially, 20 semi-structured interviews were conducted to ascertain the need to implement a nexus for construction management within the construction domain. The interview questions were designed to understand the necessity of automating activities suitable for NLP integration and identify the nodes of the nexus. These experts were purposively selected to ensure a balanced representation across construction, digitalisation and AI industries. Table 2 provides an overview of the experts' profiles in the study.

Table 2: Profile of the experts

Nr	Designation/ Field	Experience (Years)	Country
R1	Director - Consultant QS	22	Sri Lanka
R2	Director – Contractor QS	18	Sri Lanka
R3	Engineer - Transportation Sector	18	Australia
R4	LLM Developer for Procurement	10	Australia
R5	NLP model developer for Translations	8	USA
R6	Consultant QS/5D BIM Agent	32	Sri Lanka
R7	Professor – Construction Law	20	UK
R8	Construction Contract Manager	24	UK
R9	Construction Data Analyst	8	UK
R10	Construction Project Manager	18	Sri Lanka
R11	Expert Witness in Delay Analysis	32	Sri Lanka

Nr	Designation/ Field	Experience (Years)	Country
R12	Development Manager	15	UK
R13	Contract Manager	10	Nigeria
R14	Construction Lawyer	9	UK
R15	Construction Automation Professor	20+	Australia
R16	Data Science Professor	30+	Germany
R17	Commercial Manager	15+	UK
R18	Claims Specialist	27	UAE
R19	Contract Specialist	24	UAE
R20	Chief Executive Officer/ Architect	12	UK

Following the analysis of the requirements outlined by these experts, the need to establish an NLP-PCMN was substantiated. Subsequently, a focus group was convened involving experts R4, R5, R9, and R16, proficient in AI model development, to formulate the blueprint of an NLP-PCMN. Qualitative data obtained from the discussions were subjected to content analysis, with the software NVivo 12 software facilitating systematic analysis and interpretation of the data.

4. RESEARCH FINDINGS AND ANALYSIS

4.1 IMPERATIVE FOR IMPLEMENTING AN NLP-PCMN

The importance of implementing an NLP-PCMN was investigated through expert interviews. A consensus among all respondents emphasised the criticality of effective construction management, necessitating the storage and management of a vast repository of indexed textual data within the construction domain. R1 interpreted this by stating that "*a human can't store all the information and access the information all together.*" R2, R8, R10 and R13 confirmed this by stating that this complexity can overwhelm humans, especially when sources precede the other. However, as explained by R5, AI is capable of "*Long Short-Term Memory that can access a large amount of data for a short period.*" Supporting this notion, nearly all respondents suggested that automation addresses this challenge by alleviating reliance on human memory, thereby ensuring that critical points are not overlooked. However, R14 presented a contrasting viewpoint, arguing that human interaction remains indispensable as not all information may be pertinent to resolving a legal matter. This contention was contested by R4, who proposed a model capable of training on identifying precedence.

Moreover, R7, R8, R9, R17, and R20 highlighted that many professionals engaged in contract management lack proficiency in legal affairs, often necessitating reliance on legal experts for guidance. This dependency can result in delays, costs and inefficiencies. Nevertheless, integrating legal considerations into automation fosters self-empowerment among professionals and minimises inefficiencies. Furthermore, R11 and R14 emphasised the significance of staying current with case laws, albeit interpreting them amidst evolving conditions and constantly changing planning regulations that can pose challenges. Additionally, adherence to various building codes, which may vary by location, is vital. As suggested by R20, automation streamlines this process by centralising and organising information, thereby enhancing accessibility.

4.2 STAKEHOLDER EMPOWERMENT THROUGH NLP-PCMN

The expert interviews identified all types of users that will benefit from a proposed NLP-PCMN and particular use cases of the nexus. Project managers at the site were referenced most by experts for the professionals who would be empowered the most. Furthermore, using this nexus as a legal handbook was predicted to be a use case for all its potential users. R1 described it: "*Although I might use books normally, I will use this in meetings because turning pages is not nice.*" Furthermore, R1, R12, R13, R15 and R17 addressed that the chatbot will provide a significant quality of life for entry-level professionals in the industry. Furthermore, 80% of the experts mentioned that clients will benefit from this innovation as they can "*stay in touch and understand the legal considerations of their project.*"

4.3 NODES OF THE NATURAL LANGUAGE-POWERED CONTRACT MANAGEMENT NEXUS

An NLP-PCMN ideally should provide a comprehensive solution to tackle diverse legal and regulatory compliance challenges within the construction industry. Through collaboration with experts, the focus group discerned that the main features are paramount: 'Contract Administration', 'Dispute Resolution and Litigation', 'Planning Code Compliance', 'Procurement Guidelines', 'Project Management' and 'Correspondence Analysis'. To effectively support these features, the nexus must be underpinned by vector databases capable of storing essential legal and regulatory data. Furthermore, a blueprint that seamlessly integrates suitable vector databases is paramount to addressing the concerns raised by R4 and R14 in assessing the need for an NLP-PCMN.

The primary requirement for 'Contract Administration' is enabling users to access pertinent contractual provisions governing specific scenarios, as R1, R18, and R19 emphasised. This process involves analysing contract documents, including clauses, legal requirements and particular conditions. Similarly, incorporating a case law library involves creating a classification system, metadata tagging and indexing. This integration facilitates informed decision-making in the contract administration by consolidating all required information. A case law library with databases for contract law provisions and dispute resolution procedures vector database will formulate a 'Dispute Resolution and Litigation' feature as proposed by R11.

Procurement guidelines are often subject to updates by government agencies, and they present a challenge in tracking and incorporating the latest amendments. As highlighted by R2, the ideal solution should furnish users with answers based on the most recent or recent past guidelines, identifying and integrating the latest changes. A similar mechanism can indeed be implemented for 'Correspondence Analysis' within the NLP-PCMN framework. Various records such as letters, requests for information documents, meeting minutes, progress reports, and instructions can be stored systematically, akin to the approach outlined previously. This repository of documents helps project managers stay up to date on all project-related correspondence.

Furthermore, statutes and ordinances pertinent to environment law protocols, labour law provisions, and health and safety guidelines can be made accessible through vector databases, enabling project management professionals to access this critical information swiftly. Building code compliance is another crucial regulatory aspect that project managers must adhere to. This can be achieved through specialised LLMs.

As articulated by R20, planning codes exhibit variations and are often dispersed across different urban councils. The 'Planning Code Compliance' solution involves storing these guidelines in a database and referencing attributes such as building type, location, height requirements and diverse compliance constraints for building elements. Subsequently, users can query the database by providing these attributes to ascertain the appropriate constraints governing the design of a building on a vacant plot. This approach streamlines the process of accessing and navigating planning codes during the design phase of construction projects.

4.4 IMPLEMENTATION BLUEPRINT OF THE NLP - PCMN

To implement the NLP-PCMN, in the focus group stage, the R4, R5, R9 and R16 concluded the following blueprint for its architecture. The dash lines represent the vector databases that provide sources for the specific LLM, such as 'Contract Administration'. All those primary services are connected, representing a nexus that provides an all-in-one solution for legal and regulatory compliance management, as presented in Figure 2.

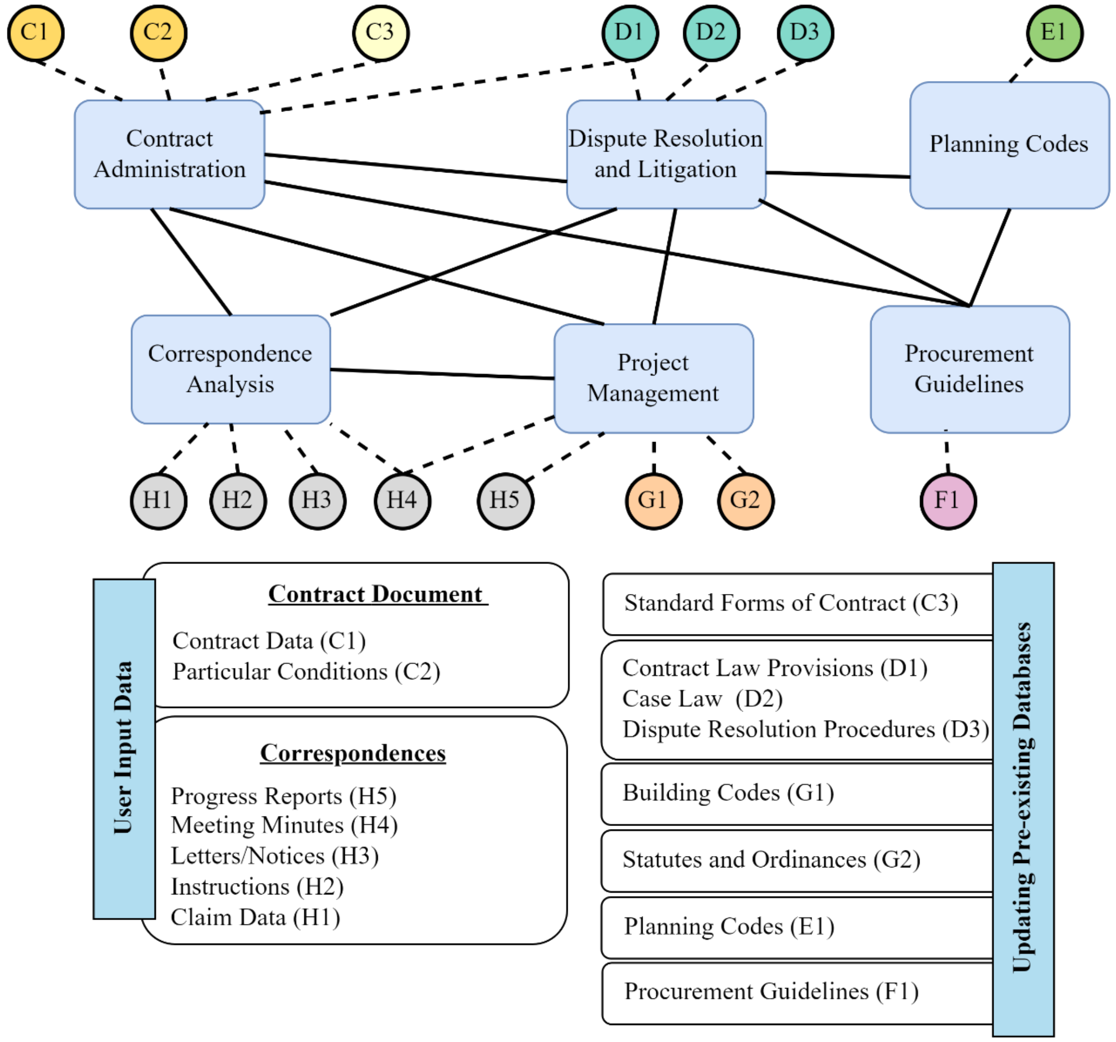

Figure 2: Blueprint for the NLP-PCMN (Source: Developed by authors)

A multi-step approach should be designed for 'Contract Administration', where a user input and pre-existing database are used. The process starts by analysing the contract document and identifying clauses, particular conditions, and contract data through keyword detection and a domain-specific LLM. Then, the document is segmented and converted into vector embeddings for efficient retrieval. When a user asks a question, the system utilises an information retrieval algorithm to extract relevant vectors. The retrieved vectors are then fed into an LLM specifically trained for contract interpretations, generating a comprehensive response that includes the answer, related clauses and document references. Figure 3 illustrates the above process.

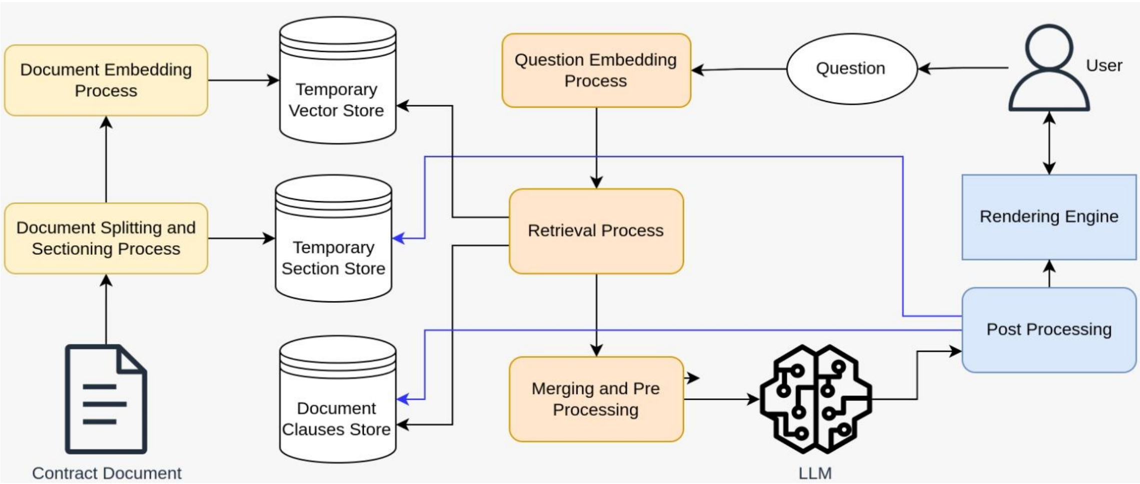

Figure 3: Synthesising user input data with pre-existing databases (Source: Developed by authors)

For 'Correspondence Analysis' and 'Procurement Guidelines Compliance', it is vital that up-to-date information is updated within the database. The following paragraphs describe the implementation of such a system as proposed by R4. An information retrieval algorithm segments the guidelines into individual clauses and extracts crucial information such as the procurement guideline and the year. Incorporating new guidelines requires adding them as new rows in the vector database. When users pose questions, the query undergoes embedding using the same mechanism as Figure 3. Following the retrieval of relevant clauses, a prompt is formulated utilising the retrieved information and the user's question. A check is conducted to identify any duplicate clauses for the same clause number, if applicable, indicating them as "Previous" and "New" within the prompt. Subsequently, a comprehensive answer is generated.

After analysing the literature alongside this study's findings, several key points emerge. The models discussed address specific tasks within contract management or the legal landscape. In contrast, the proposed architecture aims to integrate all necessary functions for managing legal and contractual aspects in the construction industry into one comprehensive nexus. This holistic approach ensures that various tasks are handled within a single, integrated system. Nevertheless, this study pioneers a consolidated NLP-PCMN architecture. Moreover, this architecture leverages the capabilities of LLMs, representing the forefront of NLP technology (Cambria & White, 2014). By using LLMs, the system can provide more accurate and context-aware insights, making it a powerful tool for construction's legal and contract landscape.

Furthermore, the NLP-powered models that were introduced in the literature are tailored to particular organisations. As concluded by R1, a publicly accessible NLP-PCMN can democratise access to legal information. As Mitchell and Mancoridis (2006) highlighted this architecture offers significant advantages. This enables individuals and businesses to navigate the law without expensive legal consultations (Mitchell & Mancoridis, 2006).

5. CONCLUSIONS AND RECOMMENDATIONS

This study was directed to develop a blueprint for implementing an NLP-PCMN in the construction industry. The experts that were interviewed all highlighted the potential use cases and the need to implement an NLP-PCMN. One of the key considerations was the complexity and vast amount of textual data in construction. Therefore, effective construction management necessitates efficient storage and management of vast textual data, a task beyond human capacity alone. Automation through NLP-PCMN not only addresses this challenge by alleviating reliance on human memory but also empowers stakeholders across various levels of expertise. Stakeholder empowerment through NLP-PCMN extends to project managers, entry-level professionals, legal experts, and clients alike. By serving as a legal handbook accessible during meetings and enhancing the quality of life for industry newcomers, the NLP-PCMN promises to revolutionise workflows and decision-making processes.

The implementation blueprint for NLP-PCMN, carefully crafted by expert focus groups, underscores six significant features required to realise its full potential i.e. (i) ‘Contract Administration’, (ii) ‘Dispute Resolution and Litigation’, (iii) ‘Planning Code Compliance’, (iv) ‘Procurement Guidelines’, (v) ‘Project Management’ and (vi) ‘Correspondence Analysis’. The nexus is powered by the integration of domain-specific vector databases. Those databases are formed by segmenting different documents in the construction industry through keyword detection and metadata tagging and then converted into vector embeddings for efficient retrieval. Furthermore, by facilitating the constantly updating legal framework, the nexus model minimises the risk of non-conformance.

Overall, by analysing all these findings, the NLP-PCMN represents a significant advancement in the construction industry, offering a comprehensive solution that addresses the intricate challenges faced by professionals in the legal and regulatory landscape. Furthermore, it could lay the foundation for enhanced collaboration, compliance and decision-making in the future of the construction industry.

To enhance the effectiveness and relevance of the NLP-PCMN, several key recommendations are proposed. Firstly, the research findings emphasise the need for NLP-powered tools in the construction industry. Therefore, it is recommended that construction industry practitioners invest in integrating NLP-PCMN systems into their existing management frameworks. Ideally, government agencies and professional institutes should develop a publicly accessible NLP-PCMN system to ensure reliability and widespread use. For research purposes, user-centric and publicly available NLP models should be developed. Developers should use the requirements of the construction industry, as identified in this study, to create effective solutions for the legal and contractual management landscape in the construction industry.

This research contributes to academia by filling the gap for a consolidated architecture of an NLP-PCMN. Additionally, it pioneers the concept of a consolidated NLP solution to the construction industry as an NLP-PCMN. It is also important to acknowledge the limitations of this study. This research considers the technological advancements up to April 2024. Grounded in interpretivism, the study acknowledges the subjectivity inherent in qualitative research, particularly when combining the knowledge from NLP developers and construction practitioners. These limitations highlight the need for future studies to consider broader samples and employ rigorous methods for interpreting qualitative data. Furthermore, Future research should include a case study using a developed model to identify potential time and cost savings as well as studies on technology adoption of NLP-powered models in the construction industry.

6. REFERENCES

360iResearch (2024). *Global AI in construction market by offering (services, solutions), deployment type (cloud, on-premises), organisation size, application, industry type: Forecast 2024-2030* (Report No. 5789100). United States. G. M. I. Inc. https://www.researchandmarkets.com/report/construction#:~:text=The%20AI%20in%20Construction%20Market,USD%203%2C226.92%20million%20by%202030

Babatunde, I., Ezirim, A., & Bid, O. (2023, September 3-9). *Enhancing contract management through Natural Language Processing(NLP): A case study of three African countries* [Paper presentation]. 5th Deep Learning Indaba, Accra, Ghana. https://openreview.net/forum?id=DUrRx78LEw

Beach, T. H., Hippolyte, J. L., & Rezgui, Y. (2020). Towards the adoption of automated regulatory compliance checking in the built environment. *Automation in Construction*, *118*(12). https://doi.org/10.1016/j.autcon.2020.103285

Beach, T. H., Rezgui, Y., Li, H., & Kasim, T. (2015). A rule-based semantic approach for automated regulatory compliance in the construction sector. *Expert Systems with Applications*, *42*(12), 5219-5231. https://doi.org/10.1016/j.eswa.2015.02.029

Cambria, E., & White, B. (2014). Jumping NLP curves: A review of natural language processing research. *IEEE Computational Intelligence Magazine*, *9*(2), 48-57. https://doi.org/10.1109/MCI.2014.2307227

Hand, D. J., Christen, P., & Kirielle, N. (2021). F*: An interpretable transformation of the F-measure. *Machine Learning*, *110*(3), 451-456. https://doi.org/10.1007/s10994-021-05964-1

Hassan, F., & Le, T. (2020). Automated requirements identification from construction contract documents using natural language processing. *Journal of Legal Affairs and Dispute Resolution in Engineering and Construction*, *12*(2). https://doi.org/10.1061/(ASCE)LA.1943-4170.0000379

Jallan, Y., Brogan, E., Ashuri, B., & Clevenger, C. M. (2019). Application of natural language processing and text mining to identify patterns in construction-defect litigation cases. *Journal of Legal Affairs and Dispute Resolution in Engineering and Construction*, *11*(4). https://doi.org/10.1061/(ASCE)LA.1943-4170.0000308

Kelley, G. (2012). *Construction law: An introduction for engineers, architects, and contractors*. Wiley & Sons. http://ndl.ethernet.edu.et/bitstream/123456789/13047/1/116.pdf

Khurana, D., Koli, A., Khatter, K., & Singh, S. (2023). Natural language processing: State of the art, current trends and challenges. *Multimedia Tools and Applications*, *82*(3), 3713-3744. https://doi.org/10.1007/s11042-022-13428-4

Lauriola, I., Lavelli, A., & Aiolli, F. (2022). An introduction to deep learning in natural language processing: Models, techniques, and tools. *Neurocomputing*, *470*, 443-456. https://doi.org/10.1016/j.neucom.2021.05.103

Lee, J., Ham, Y., Yi, J.-S., & Son, J. (2020). Effective risk positioning through automated identification of missing contract conditions from the contractor's perspective based on FIDIC contract cases. *Journal of Management in Engineering*, *36*(3). https://doi.org/10.1061/(ASCE)ME.1943-5479.0000757

Lee, J.-K. et al. (2023). High-level implementable methods for automated building code compliance checking. *Developments in the Built Environment*, *15*(12). https://doi.org/10.1016/j.dibe.2023.100174

Lee, J., Yi, J.-S., & Son, J. (2019). Development of automatic-extraction model of poisonous clauses in international construction contracts using rule-based NLP. *Journal of Computing in Civil Engineering*, *33*(3). https://doi.org/10.1061/(ASCE)CP.1943-5487.0000807

Madan, R., & Ashok, M. (2023). AI adoption and diffusion in public administration: A systematic literature review and future research agenda. *Government Information Quarterly*, *40*(1). https://doi.org/10.1016/j.giq.2023.101881

Marzouk, M., & Enaba, M. (2019). Text analytics to analyse and monitor construction project contract and correspondence. *Automation in Construction*, *98*, 265-274. https://doi.org/10.1016/j.autcon.2018.11.018

Mitchell, B. S., & Mancoridis, S. (2006). On the automatic modularisation of software systems using the bunch tool. *IEEE Transactions on Software Engineering*, *32*(3), 193-208. https://doi.org/10.1109/TSE.2006.31

Moon, S., Lee, G., & Chi, S. (2022). Automated system for construction specification review using natural language processing. *Advanced Engineering Informatics*, *51*. https://doi.org/10.1016/j.aei.2021.101495

Murdoch, J., & Hughes, W. (2002). *Construction contracts: Law and management* (5th ed.). Routledge. https://www.nibmehub.com/opac-service/pdf/read/Construction%20Contracts%20Law%20and%20Management%205th%20Edition.pdf

Nadkarni, P. M., Ohno-Machado, L., & Chapman, W. W. (2011). Natural language processing: An introduction. *Journal of the American Medical Informatics Association*, *18*(5), 544-551. https://doi.org/10.1136/amiajnl-2011-000464

Padhy, J., Jagannathan, M., & Delhi, V. (2021). Application of natural language processing to automatically identify exculpatory clauses in construction contracts. *Journal of Legal Affairs and Dispute Resolution in Engineering and Construction*, *13*(4). https://doi.org/10.1061/(ASCE)LA.1943-4170.0000505

Parikh, P. M., Shah, D. M., & Parikh, K. P. (2023). Judge Juan Manuel Padilla Garcia, ChatGPT, and a controversial medicolegal milestone. *Indian Journal of Medical Sciences*, *75*(1), 3-8. https://doi.org/10.25259/IJMS_31_2023

Potrac, P., Jones, R. L., & Nelson, L. (2014). Interpretivism. In L. Nelson, R. Groom & P. Potrac (Eds.), *Research Methods in Sports Coaching* (pp. 31-41). Routledge. https://doi.org/10.4324/9780203797549-4

Saunders, B., Sim, J., Kingstone, T., Baker, S., Waterfield, J., Bartlam, B., Burroughs, H., & Jinks, C. (2018). Saturation in qualitative research: Exploring its conceptualisation and operationalisation. *Quality & Quantity*, *52*(4), 1893-1907. https://doi.org/10.1007/s11135-017-0574-8

Shaikh, A., & Gohar, S. (2024, May 30-31). *Study on use of text based AI Tool in construction industry Contracts* [Paper Presentation]. 3rd IBA SBS International Conference 2024, Karachi, Pakistan. https://ir.iba.edu.pk/sbsic/2024/program/59

Soibelman, L., Wu, J., Caldas, C., Brilakis, I., & Lin, K.Y. (2008). Management and analysis of unstructured construction data types. *Advanced Engineering Informatics*, *22*(1), 15-27. https://doi.org/10.1016/j.aei.2007.08.011

Stone, P., Brooks, R., Brynjolfsson, E., Calo, R., Etzioni, O., Hager, G., Hirschberg, J., Kalyanakrishnan, S., Kamar, E., & Kraus, S. (2021). *Artificial intelligence and life in 2030: The one hundred year study on artificial intelligence*. United States. University of Stanford. https://doi.org/10.48550/arXiv.2211.06318

Szewc, T. (2022). The impact of legal regulations on investment project management in construction. *Organisation and Management: Scientific Quarterly*, *1*(5), 145-154. https://doi.org/10.29119/1899-6116.2021.57.10

Wu, C., Li, X., Guo, Y., Wang, J., Ren, Z., Wang, M., & Yang, Z. (2022). Natural language processing for smart construction: Current status and future directions. *Automation in Construction*, *134*. https://doi.org/10.1016/j.autcon.2021.104059

Yan, H., Yang, N., Peng, Y., & Ren, Y. (2020). Data mining in the construction industry: Present status, opportunities, and future trends. *Automation in Construction*, *119*(1). https://doi.org/10.1016/j.autcon.2020.103331

Zhang, D., Mishra, S., Brynjolfsson, E., Etchemendy, J., Ganguli, D., Grosz, B., Lyons, T., Manyika, J., Juan, Sellitto, M., Shoham, Y., Clark, J., & Perrault, R. (2022). *The AI index 2022 annual report*. Stanford University. https://doi.org/10.48550/arXiv.2205.03468