Kalpa Publications in Computing
Volume 22, 2025, Pages 570–577
Proceedings of The Sixth International Conference on Civil and Building Engineering Informatics

Heritage Building Information Management and Intelligent Querying by Multimodal Large Language Models and Knowledge Graph

Jiaying ZHANG[1], Jeff Chak Fu CHAN[2], Ziyu ZHAO[3], Jack C P CHENG[4*]
[1] Postdoctoral Fellow, Department of Civil and Environmental Engineering, Hong Kong University of Science and Technology, Hong Kong, China. Email: jzhangfh@connect.ust.hk
[2]M.Phil. Student, Department of Civil and Environmental Engineering, Hong Kong University of Science and Technology, Hong Kong, China. Email: cfchanay@connect.ust.hk
[3]Ph.D, Department of computer science and software engineering, the University of Western Australia, Australia. Email: zhaoziyu929@126.com
*Corresponding author. Professor, Department of Civil and Environmental Engineering, The Hong Kong University of Science and Technology, Hong Kong, China. 999077. Email: cejcheng@ust.hk.

Abstract

Heritage buildings face challenges in documentation due to inconsistent records and complex data from historical documents, archaeological surveys, and materials. Traditionally, converting unstructured data into structured formats required significant expert effort. The advent of large language models (LLMs) has transformed heritage research by enabling the creation and maintenance of knowledge graphs. These graphs integrate diverse data sources, facilitating the preservation and study of heritage buildings. LLMs help extract and organize unstructured data, improving knowledge graph accuracy and consistency. This research proposes a comprehensive framework that integrates multimodal data, including text, images, and videos, into a unified knowledge graph. The framework employs LLMs for extracting information from textual data, the CLIP model for aligning images with corresponding text, and keyword searches for processing video content. The resulting knowledge graph is stored in a Neo4j graph database, providing an interactive platform for users to query and explore detailed information about heritage buildings. This approach not only supports academic research but also contributes to practical applications in cultural heritage conservation, enabling more efficient access to valuable information and enhancing preservation efforts. The proposed method was validated in European 'Gothic' and 'Gothic Revival' architecture by comparing the relationships between components.

1 Introduction

Heritage buildings have gone through different uses and eras, as well as multiple periods of changes and advancements in technologies and regulations (Cheng et al., 2024). Information on these buildings comes from historical documents, archaeological surveys, building materials, academic research, and more, often involving inconsistent and complex records. For example, structural design, construction techniques, and materials recorded in ancient architectural documents may differ from the current state of the buildings. Traditionally, converting unstructured data into structured formats required significant human effort, particularly for heritage buildings, where domain experts had to process the data (AP News, 2024). Previously, acquiring knowledge about heritage buildings required extensive learning and literature review (AP News, 2024). Khalil et al (2021) indicates that the documentation of cultural heritage faces additional challenges, primarily due to the varying sizes of the documented objects and the differing requirements for quality and resolution. The purpose of documentation is to make information accessible to other users. Challenges include stability, surface damage, variations in scale and precision requirements, time constraints, and the involvement of multi-sensor and multi-resolution data. The documentation of cultural heritage is not an end but aims to provide information to others. Patias et al (2006) demonstrated that main difficulties in generating traditional knowledge graphs include the complexity of data acquisition and integration, challenges in information extraction and semantic understanding, technical issues in knowledge representation and storage, and the accuracy of relationship extraction and entity linking. Additionally, maintaining the dynamic updating and reasoning capabilities of knowledge graphs poses demands on their application effectiveness and reliability.

The rise of large language models (LLMs) has brought significant changes to how we study and disseminate knowledge of ancient architecture. Previously, learning about heritage buildings required extensive literature reviews. Documentation of cultural heritage faces challenges, including varying object sizes, damage, time constraints, and multi-sensor data. The purpose of documentation is to provide accessible information to users, though challenges remain in maintaining accuracy and consistency. LLMs, through deep learning and pre-training, have addressed many natural languages processing challenges, including language understanding, multitasking, and cross-domain applications. They offer new opportunities for constructing, updating, and reasoning through knowledge graphs, which can enhance the management and application of historical building data. A knowledge graph integrates various data sources such as historical literature and multimedia resources, creating a comprehensive database for preservation and restoration efforts. sing LLMs to build a knowledge graph for historical building components allows for effective extraction and organization of information from unstructured data. This process involves data collection, preprocessing, information extraction, and continuous updating. The resulting knowledge graph, stored in a graph database, enables efficient querying and reasoning, supporting the research and preservation of historical buildings. The graph also aids cultural heritage protection, providing valuable insights for tourism, smart guide systems, and personalized travel experiences. By integrating various data sources and maintaining consistency, the system ensures high-quality information that supports both academic and practical applications in building restoration.

2 Proposed method

In this research, we present a framework, which integrates multimodal data to enhance structured representations. Firstly, we use LLM via prompt engineering to extract informative triplets. Meanwhile, the CLIP model is applied to align images with given text input, as well as video data via key word search. It constructs a knowledge graph which is stored in neo4j graph database and can be interacted with via an interactive LLM-based chat agent. Ultimately, the goal is to develop an LLM-based chatbot interface that leverages this enriched knowledge graph to provide users with sophisticated query capabilities. The proposed method is shown in Figure 1.

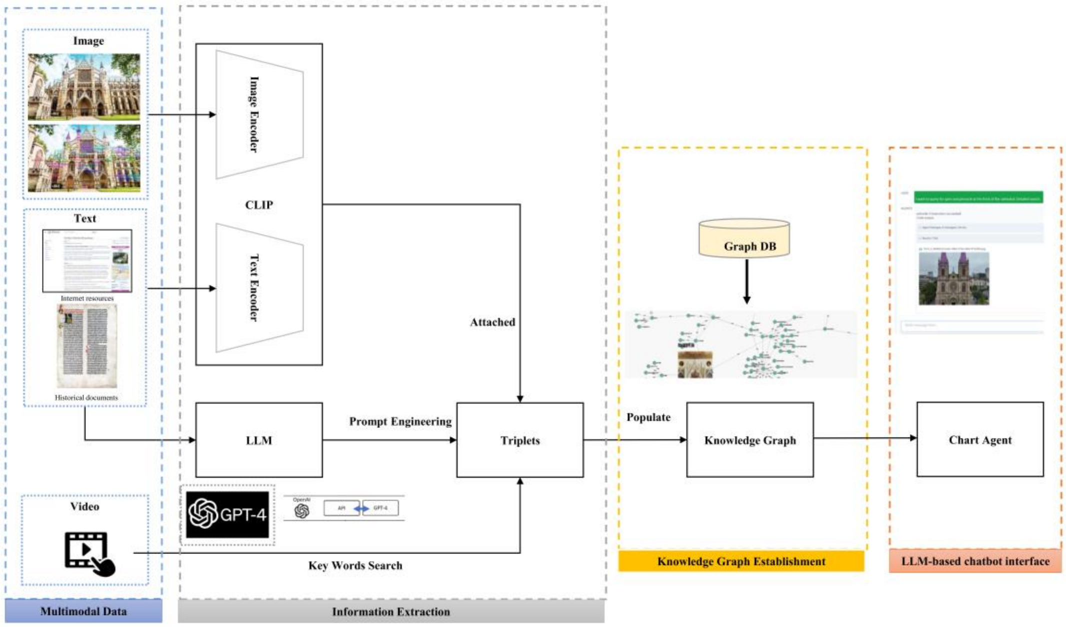


Figure 1: The proposed method

2.1 Database Generation and knowledge graph

A knowledge graph can clearly express the hierarchical relationship of architectural heritage knowledge (Fan et al., 2021; Zhang et al., 2022). Building a knowledge graph of Gothic architecture using large language models starts with collecting textual data about Gothic architecture, including its stylistic features, origins, representative buildings, development history, and especially semantic information about architectural structures. Secondly, the collected textual data is processed and analyzed to extract key information and knowledge points. Entities and relationships are established in the knowledge graph, organizing the various aspects of Gothic architecture into structured knowledge graph data and identifying their intrinsic connections. Conventional knowledge graph construction requires a large amount of textual data for training. This study proposes a large language self-learning model that can accurately present the target building with just a small set of data samples Thirdly, as the cultural significance and symbolic features behind ancient architecture have a long history and evolve over time, with the emergence of new content and the construction of new buildings imitating ancient styles, it is necessary to continue updating and optimizing the knowledge graph data to ensure timeliness and accuracy. Combining user feedback and requirements, continuously improving the retrieval efficiency of the knowledge graph, and enhancing user experience.

2.2 Querying ability of the LLM-based assistant

Semantic recognition of heritage building components can be very challenging for non-experts, but AI large language models (LLMs) provide a dialogue interaction agent for building digital models of heritage structures. The digital twin agent can engage in conversation with users, who only need to input vague query instructions or capture images of building components, and the agent will provide accurate and relevant information, as well as query the building's background and historical context. Through interaction with the AI large language model, users can quickly retrieve information about the historical background, architectural style, and cultural significance of heritage buildings, greatly improving the documentation and querying capabilities of components in heritage digital twins. The query capability of LLMs is a core advantage in data management, information retrieval, and decision support, and it manifests in several key ways: first, LLMs support natural language queries, allowing users to express their queries in everyday language, thus lowering the technical barrier and accommodating various query expressions. Second, LLMs can process complex multi-condition queries and fuzzy queries, helping users find data that is not an exact match. Third, LLMs provide intelligent recommendations and auto-completion features, offering real-time query suggestions and automatically completing queries to improve efficiency. Fourth, LLMs possess strong contextual understanding, comprehending the semantics of queries rather than just keyword matching, and can retain context throughout a conversation to support continuous queries. Finally, LLMs support multilingual queries, making them useful in cross-lingual environments for data query and management. In practical applications, users can query the historical background, restoration records, and current status of heritage buildings using natural language, for example, "query the restoration history of Notre-Dame Cathedral" or "find all 18th-century wooden structures," which greatly facilitates the querying and management of heritage building data.

3 Illustrative example

Using the extremely famous Gothic architectural style as the training theme, this research has sorted out the names, characteristics, and interrelationships of components in European 'Gothic' and 'Gothic Revival' architecture. We have organized 20 illustrated documents as training materials for entity extraction, highlighting the differences between the two styles. The comparison of Gothic and Gothic Revival are summarized in Table, examples of the relationships between components of Gothic architecture extracted from the text are listed in Table.

Name of Components	Gothic	Gothic Revival
Pointed Arch	Basic load-bearing element, sharp and towering	Serve as a decorative element, with more varied forms
Ribbed Vault	Light and stable, supporting large spaces	Be simplified or use modern materials and techniques
Flying Buttress	Important load-bearing structure, enhancing stability	Be smaller or absent, depending on design requirements
Rose Window	Iconic decorative element, intricate and ornate	Be more simplified or have more complex decorations
Building Materials	Primarily stone and glass	Incorporate a variety of materials
Overall Style	Mysterious, solemn, and soaring	Retro, diversified, and possibly incorporating innovative designs

Table 2: The comparison of Gothic and Gothic Revival

Entity 1	Entity Type 1	Relationship	Entity 2	Entity Type 2
Pointed arches	component	Bear	Buttresses	component
Buttresses	component	prop up	Fly buttresses	component
Stained glass	component	Installed on	walls	Architectural parts
Beam columns	component	prop up	vault	component
Spire	component	connect	The main body of the building	building
Sharp coupons	component	cooperate	Clover arch	Decorative components
Clover arch	Decorative components	ornament	Four-leaf windows	Decorative components
Pointed arches	component	belong	Gothic architecture	Construction period
Spire	component	belong	Gothic Revival	Construction period
Fly buttresses	component	prop up	Nave vault	Architectural parts
Ribbed vaults	component	prop up	Chapel space	Architectural parts
Small minarets	Decorative components	ornament	Fly buttresses	component

Table 3: Examples of the relationships between components of Gothic architecture extracted from the text

3.1 Database Generation and knowledge graph

Gothic architectural elements not only have specific names with literary connotations but also interact with one another and serve distinct functions. For example, the flying buttress is an external structure used in Gothic architecture to support the vaulted ceilings. Through a series of diagonal supports, it transfers the weight of the roof to the foundation, reducing the pressure on the interior columns. The introduction of flying buttresses allowed Gothic cathedrals to have taller and more open spaces, while also giving architects more freedom to incorporate elements like pointed arches and rose windows into their designs. Ribs are one of the most iconic features of Gothic architecture. They consist of a series of arched structures that distribute the weight to the walls and columns, allowing the building to have higher spaces and lighter structures. The ribs are typically in the shape of pointed arches, which are not only aesthetically pleasing but also capable of supporting larger spans, making the interior of the cathedral more spacious and well-lit. The relationship of building components of heritage building with Gothic and Gothic Revival style are investigated in the knowledge graph (Figure 2a-c).

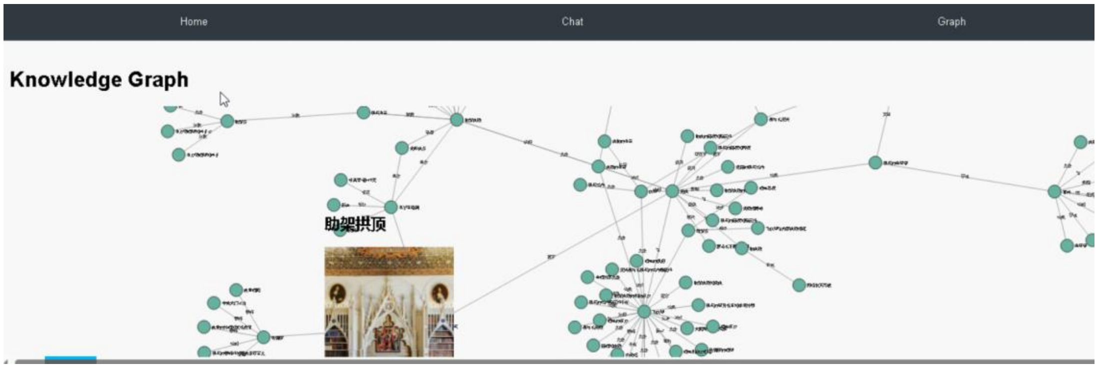


Figure 2a: Knowledge graph of building components

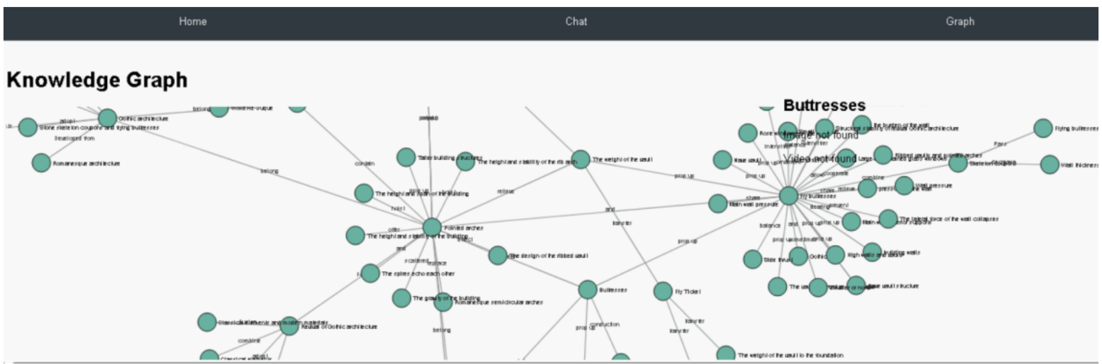


Figure 2b: Knowledge graph of building components

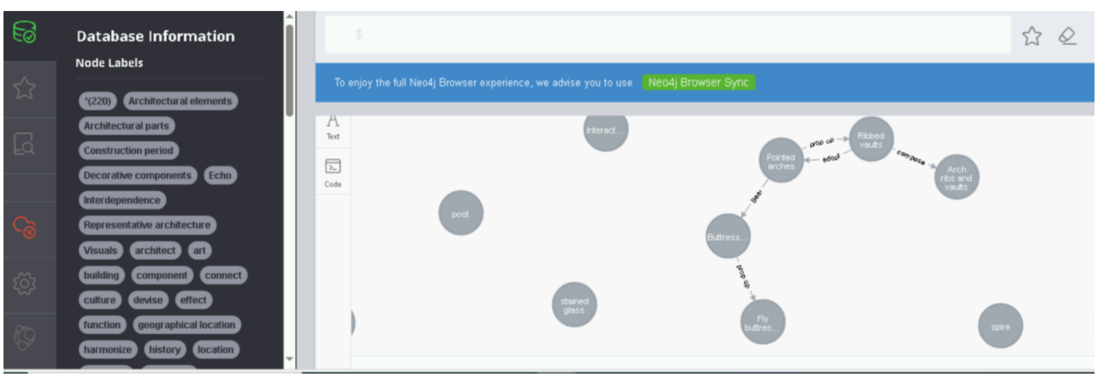
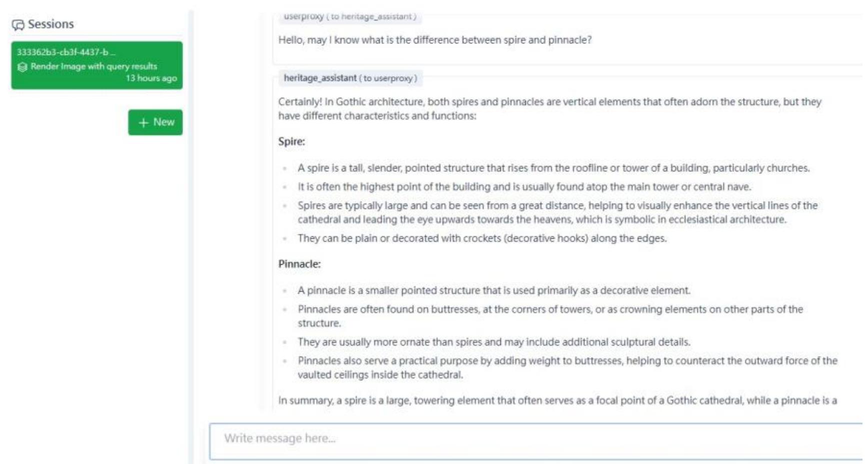


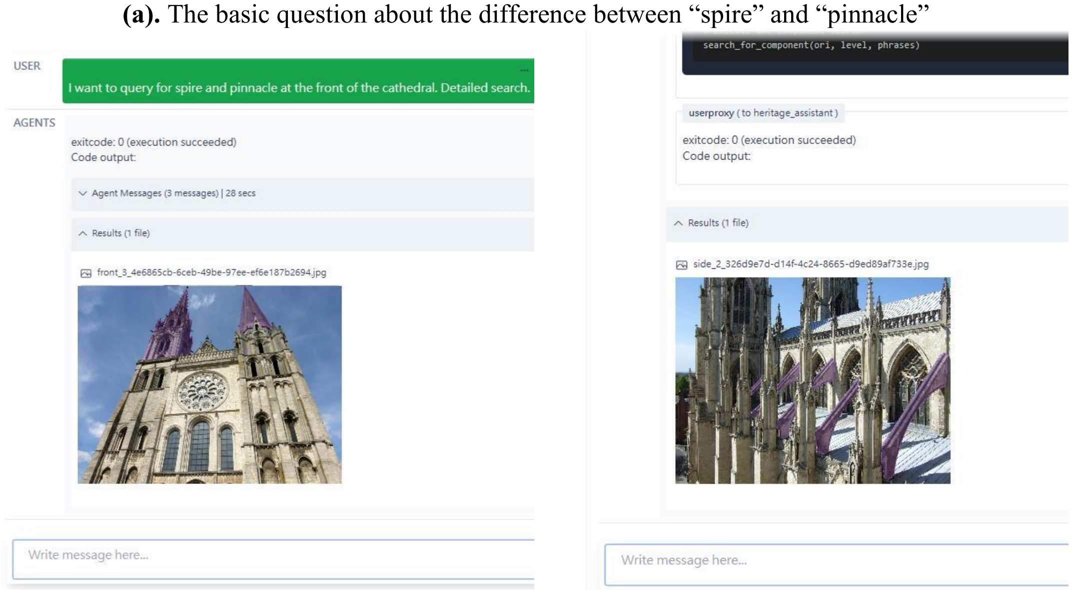
Figure 2c. Knowledge graph of building components

3.2 LLM-based Intelligent Querying

The visualization of querying related to components within the LLM-based chatbot interface of the application was demonstrated in Figure 3. The interface was built via Autogen Studio and the LLM heritage assistant was based on GPT-4 with specific system message for instruction. We first asked the heritage assistant a basic question about the difference between “spire” and “flying buttress”, and the assistant could answer correctly as shown in Figure 3 (a). It showed that the assistant can correctly explain the meaning of both components and differentiate their characteristics and functions. The LLM assistant was provided with the details of the cosine similarity search function on language features in the backend, so that it could provide suitable input parameters to the system given the user's query. It can identify and extract the wanted view and semantic level of the model and the search keywords from user's natural language query, and pass them as input parameters to the function so that the system can execute the function correctly and return an annotated image with query result, as shown in Figure 3 (b). Not only can the LLM assistant extract clear search keywords from the user's query, but it can also transform vague words in the query like "big round window" and "flying arched column" into standard component names like "rose window" and "flying buttress" to conduct the similarity search. Through a simple demonstration application, it is shown that an LLM-based assistant could improve the documentation and querying ability of components in the heritage, facilitating user-friendly navigation and information retrieval of the heritage.



(a). The basic question about the difference between "spire" and "pinnacle"



(b). Query results of the "spire" and "flying buttress" at the front and side views

Figure 3: Question and answering in the chatbot interface regarding the previous image

4 Conclusions

This research proposes a framework that integrates multimodal data to enhance the structured representation of heritage building knowledge. The framework leverages Large Language Models (LLMs) with prompt engineering to extract informative triplets from textual data. In parallel, the CLIP model is employed to align images with corresponding text inputs, and video data is processed using keyword searches to further enrich the dataset. This multimodal data is then used to construct a comprehensive knowledge graph, which is stored in a Neo4j graph database. The knowledge graph is designed to be interactive, with users able to query and explore the data through an LLM-based chatbot interface. This chatbot allows users to make sophisticated queries and retrieve detailed information about historical buildings, supporting both academic research and practical applications in cultural heritage conservation. The goal of this framework is to enable seamless access to enriched knowledge about heritage buildings, facilitating advanced data exploration and enhancing user experience.

5 References

Cheng, J.C., Zhang, J., Kwok, H.H., Tong, J.C. (2024). Thermal performance improvement for residential heritage building preservation based on digital twins. Journal of Building Engineering 82, 108283.

Fan, Q., Shi, Z., On, G. (2021). Construction of knowledge graph of Intangible Cultural Heritage, Libr. Forum. pp. 100-109.

Khalil, A., Stravoravdis, S., Backes, D. (2021). Categorisation of building data in the digital documentation of heritage buildings. Applied Geomatics 13, 29-54.

AP News, (2024). Next-Gen Digital Twin and AI Breathe New Life into Architectural Heritage Preservation: Innovative Research Released. https://apnews.com/press-release/ein-presswire-newsmatics/cultural-preservation-hong-kong-68b0a82c70d9103ff8d563dfc6fd7471.

Patias, P. (2006). Cultural heritage documentation. International Summer School Digital Recording and 3D Modeling, Aghios Nikolaos, Creta, Grecia, Abril, 24-29.

Zhang, X., Zhi, Y., Xu, J., Han, L. (2022). Digital protection and utilization of architectural heritage using knowledge visualization. Buildings 12(10), 1604.