ISSN 2616-7433 Vol. 7, Issue 8: 49-57, DOI: 10.25236/FSST.2025.070808

# Construction and Application of Lingnan Architectural Culture Corpus under the Background of Generative Artificial Intelligence

**Linhong Li, Min Wang***

*Guangdong Construction Polytechnic, Guangzhou, Guangdong Province, China*

***Abstract:*** *While generative artificial intelligence is being fully integrated into architectural heritage narratives and design applications, the structured representation of Lingnan architectural culture knowledge remains lacking. This paper selects the Xiguan historical district in Guangzhou and the Ancestral Temple-Lingnan Tiandi area in Foshan as research areas, creating a multi-source corpus integrating local chronicles, planning texts, field investigations, and oral histories. An ontological framework of "time period—building type— component — craft—imagery" is presented, and unified semantic embedding is obtained through domain pre-training and graph representation. Based on this, a generative architecture integrating retrieval enhancement and lightweight fine-tuning is formed, and evaluation is conducted for three types of tasks: knowledge question answering, architectural description, and design assistance. The results show that this corpus significantly improves the recall rate of cultural entities and the consistency of Lingnan style, and also outperforms general models in terms of design efficiency and text quality, providing technical support for the digital inheritance and intelligent generation of regional architectural culture.*

***Keywords:*** *Lingnan architecture; generative artificial intelligence; cultural corpus; retrieval-enhanced generation; knowledge graph; design assistance*

## 1. Introduction

Lingnan architectural culture under digital background has new requirements for inheritance. Qing and Wang [1] proposed a multimodal intangible cultural heritage narrative framework, which emphasizes the complementary relationship between text, image and voice. Zhang et al. [2] analyzed the symbol system of Lingnan ancestral halls from the perspective of cultural geography. Qing and Wang [3] showed that dialect and cultural symbols are of great significance to regional identity. Zhong [4] focused on the ecological expression of traditional architecture in modern landscape and its semantic reproduction. An [5] formed an intangible cultural heritage translation corpus to highlight the transmission value of domain corpus. Du et al. [6] used knowledge graph to express the structural connection of embroidery culture. Hou et al. [7] used component-image mining to confirm the cultural analysis driven by corpus. Yazici and Ozturk [8] explored architectural language through text mining. Elgibreen et al. [9] proposed an incremental formation strategy for cultural corpus. Fan and Chen [10] used CuPe-KG to complete the association and inference of cultural resources. In summary, despite numerous achievements, a unified structure and generative adaptation mechanism are still lacking for Lingnan architectural culture corpora used in generative AI. This paper constructs a structured, semantic, and generative corpus and explores its applications in question answering, style generation, and design assistance.

## 2. Demand Analysis and Overall Corpus Architecture of Lingnan Architectural Culture

### *2.1 Application Scenarios and User Needs Analysis*

The study area selected the Xiguan Historical District in Liwan District, Guangzhou, and the Ancestral Temple-Lingnan Tiandi area in Foshan. Combining the results of multimodal narrative and spatial art analysis of Lingnan intangible cultural heritage, three core application scenarios can be summarized: scholars need to conduct knowledge-based question - and- answer sessions and implement semantic retrieval in the context of historical context and ritual vocabulary; planning and architecture professionals need to provide stylistic references for types such as arcade buildings and ancestral halls, and complete the creation of text-image prompts; cultural and tourism institutions and the public need to

ISSN 2616-7433 Vol. 7, Issue 8: 49-57, DOI: 10.25236/FSST.2025.070808

complete the automatic generation and modification of explanatory texts and educational content. The scenarios mentioned above can be extracted into four elements to form a demand vector, namely $\mathbf{d}=(T,B,G,L)$, where $T$ represents the time range, $B$ covers the type of architecture, $G$ focuses on the degree of spatial or regional segmentation, $L$ focuses on the richness of language style, and the system should ensure the same semantic alignment when performing full-text search, entity-level search, and generation tasks based on prompts.

### *2.2 Design of the Types and Hierarchical Structure of Lingnan Architectural Culture Corpus*

If we categorize the corpus types according to their source and function, we can divide them into several levels, such as historical documents, specifications and atlas descriptions, field surveys and surveying reports, conservation plans and renewal schemes, design specifications and review texts, and even oral histories and dialect annotations. These levels together encompass the complete evolutionary process from Ming and Qing ancestral halls and modern arcade buildings to contemporary renewal projects. Each level is directly compatible with subsequent knowledge graphs and vector retrieval modules through unified entity identifiers and paragraph-level annotation interfaces. On this basis, a multi-layered semantic structure of "time representative— geographical unit—building type—component system— craftsmanship— cultural imagery" is created. The research area is divided into many blocks and building clusters. Components such as gable walls, plasterwork, and arcades are encoded using discrete feature vectors, and then combined with contemporaneous and spatial indexes to create a tensor-based organizational structure.

### *2.3 Overall System Architecture and Data Flow of the Corpus*

As shown in Figure 1, the corpus adopts a four-layer system architecture: the acquisition layer performs multi-source access to PDF documents, planning archives, BIM/surveying results, and field texts; the processing layer integrates functions such as layout analysis, OCR error correction, sentence segmentation, and terminology recognition, and uses a task queue to complete incremental processing; the semantic layer converts text into vector indexes and creates a knowledge graph of building entities, components, processes, and images; and the service layer provides a unified interface for retrieval enhancement generation, analysis dashboards, and monitoring modules via API.

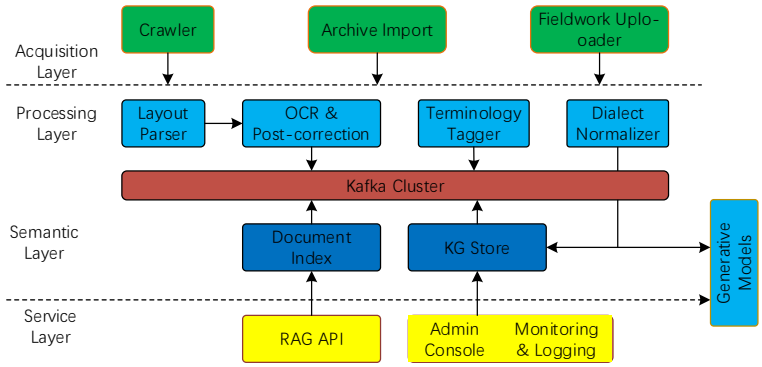


*Figure 1. Schematic diagram of the overall architecture and data flow of the Lingnan Architectural Culture Corpus System*

### *2.4 Corpus Size Planning and Coverage Index System*

To standardize the data collection process and monitor imbalances between different subdomains, a weighted coverage index $C_w$ and a domain balance index are defined $B$. Weighted coverage is used to characterize the overall completeness within the time-building type-component combination space.

$$C_w=\frac{\sum_{e\in E}\sum_{t\in T}\sum_{k\in K}w_{e,t,k}\min\left(1,\frac{N_{e,t,k}}{N^*_{e,t,k}}\right)}{\sum_{e\in E}\sum_{t\in T}\sum_{k\in K}w_{e,t,k}} \tag{1}$$

Among them, $E$ the set of eras represents the period, such as the Ming and Qing Dynasties, the

ISSN 2616-7433 Vol. 7, Issue 8: 49-57, DOI: 10.25236/FSST.2025.070808

Republic of China, and the contemporary era; the set of $T$ building types represents the building type, including ancestral halls, courtyard shops, garden residences, etc.; and the set of component types represents the component type $K$, including roof forms, firewalls, corridors, lattice windows, etc., indicating $N_{e,t,k}$ the number of texts collected under $N^*_{e,t,k}$ the current era $e$, type $t$, and component conditions. $k$ This is the text quantity target that this combination aims to achieve, $w_{e,t,k}$ and the weighting coefficients are determined according to the importance of the research and the frequency of occurrence in the research area. $C_w$ It belongs to the [0,1] interval, which reflects the overall collection progress. The closer its value is to 1, the more fully the weighted target is covered.

To avoid excessively large or small sample sizes in certain subdomains, a domain balance index is further constructed to compare the sample distribution of each corpus subdomain with a uniform distribution:

$$B = 1 - \frac{1}{2}\sum_{d=1}^{D}\left| p_d - \frac{1}{D} \right| \tag{2}$$

Here, $D$ represents the number of subdomains in the corpus, such as modern and contemporary residences, ancestral halls, arcade streets, academies/schools, etc.; $p_d = N_d / \sum_{j=1}^{D} N_j$ represents $d$ the proportion of the th subdomain in the entire corpus; $N_d$ and represents the number of texts contained in this subdomain. $B \in [0,1]$ The closer the value is to 1, the more uniform the sample distribution of each subdomain tends to be. If some subdomains are consistently underestimated, the value $B$ will decrease significantly, thus triggering dynamic adjustments to the collection priority.

Based on the above indicators, the initial target size and current data collection status of the study area can be set, and Table 1 can be generated for engineering management.

*Table 1 Statistics on the Gap between Target Corpus Size and Current Status*

| Subcorpus | Target docs | Collected docs | Avg length (tokens) | Gap ratio (%) |
|---|---|---|---|---|
| Xiguan arcaded shophouses (Guangzhou) | 2000 | 850 | 420 | 57.5 |
| Clan ancestral halls (Guangzhou & Foshan) | 1500 | 730 | 510 | 51.3 |
| Foshan Zumiao & Lingnan Tiandi blocks | 1200 | 640 | 380 | 46.7 |
| Lingnan garden residences (major cases) | 800 | 290 | 560 | 63.8 |

## 3. Corpus Acquisition, Processing, and Representation Methods (For Generative AI)

### *3.1 Design of Multi-Source Data Acquisition and Cleaning Pipeline*

The study area covers the Xiguan historical district and the Foshan Ancestral Temple-Lingnan Tiandi area. The corpus is registered in the Source Registry, with sources including local chronicles and architectural records, regulatory planning texts, survey and component records, interview transcripts, and design specifications, along with metadata such as date, coordinates, and type. The cleaning pipeline is automated end-to-end: the Layout Parser extracts layout information, OCR and a terminology dictionary jointly perform recognition and correction, the Noise Filter removes noise based on n-grams and perplexity, the Sentence Segmenter performs sentence segmentation, and the PII Anonymizer de-identifies sensitive information. Results are written to a queue in JSON-Lines format for subsequent annotation and indexing, achieving stable incremental cleaning of heterogeneous corpora.

### *3.2 Linguistic annotation and the construction of Lingnan architectural ontology*

After corpus cleaning, the linguistic annotation module performs word segmentation and part-of-speech tagging, and loads a specially designed vocabulary for Lingnan architecture, converting terms like "arcaded shophouse," "firewall gable," "veranda," and "gray plaster" into unique tokens. The "Terminology Tagger" integrates BiLSTM-CRF and rules to identify components, techniques, and imagery, and also annotates sentence-level discourse functions and text types. The ontology uses a main

ISSN 2616-7433 Vol. 7, Issue 8: 49-57, DOI: 10.25236/FSST.2025.070808

branch architecture of "architectural entity – component – technique – imagery – event," abstracting buildings, street segments, and ancestral halls into the concept of "BuildingInstance," and creating targeted associations with Component, Technique, Imagery, and Event. It utilizes RDF/Property Graph formats for storage, enabling querying and graph embedding. The corpus annotation spans are bidirectionally linked to ontology nodes via IDs, achieving enhanced retrieval and accurate restoration of cultural context.

### *3.3 Semantic Representation Methods for Text and Knowledge*

The semantic representation layer employs a multi-view strategy of "text embedding + knowledge graph embedding". The text portion is based on a domain-adapted Transformer, further pre-trained using a large-scale Lingnan architectural corpus, while additional masking weights are applied to domain terms. Its loss function is defined as:

$$\mathcal{L}_{dom} = \mathcal{L}_{LM} + \lambda_{\text{term}} \mathcal{L}_{\text{term}} \tag{3}$$

Among them, $\mathcal{L}_{dom}$ represents the overall loss of domain adaptation pre-training; $\mathcal{L}_{LM}$ refers to the loss generated by the mask prediction of the standard language model; $\mathcal{L}_{\text{term}}$ is the mask prediction loss calculated only at the location of Lingnan architectural terms; $\lambda_{\text{term}}$ and is the hyperparameter of term loss weight.

The knowledge graph component performs graph embedding learning on the ontology graph and spatiotemporal relation graph, employing GraphSAGE-style neighborhood aggregation and comparison targets.

$$\mathcal{L}_{\text{graph}} = -\sum_{v \in V} \left( \log \sigma\left(\mathbf{h}_v \mathbf{h}_{v^+}\right) + \sum_{v^-} \log \sigma\left(-\mathbf{h}_v^\top \mathbf{h}_{v^-}\right) \right) \tag{4}$$

Here $V$ is the set of nodes in the graph; $\mathbf{h}_v$ is the node $v$ embedding vector; $v^+$ is $v$ the positive sample node with semantic or structural adjacency; $v^-$ is $v$ the negative sample node with no direct relationship; $\sigma(\cdot)$ is the Sigmoid function; $\mathcal{L}_{\text{graph}}$ is the graph embedding contrastive learning loss.

The statistics of the characterization subcorpus and model parameters are shown in Table 2. All fields are numerical or Boolean variables to facilitate automatic summarization and ablation analysis.

*Table 2 Training/Representation Subcorpus and Parameter Statistics*

| Subcorpus | Samples | Total tokens | Avg sent. length | Encoder type | Dim |
|---|---|---|---|---|---|
| Historical gazetteers (Xiguan/Foshan) | 3 200 | 5.8e6 | 32 | Transformer- base | 768 |
| Planning & Regulation Texts | 2 100 | 4.1e6 | 28 | Transformer- base | 768 |
| Field survey & measurement reports | 1,450 | 2.6e6 | 35 | Transformer- base | 512 |
| Oral histories & interviews | 980 | 1.9e6 | 41 | Transformer- small | 512 |

### *3.4 Retrieval Enhancement and Fine-tuning Architecture for Generative AI*

Based on semantic representation, generative AI adopts a "retrieval enhancement + lightweight fine-tuning" architecture: after user queries are vectorized by the Query Encoder, the Vector Retriever recalls relevant candidate objects from documents and knowledge graphs, then reorders them through the Cross-Encoder, and finally integrates them into specific contextual content using the Prompt Template; on the generation side, a large model with LoRA or Prefix-Tuning is used for output, and this output is further processed by a Post-processor for terminology standardization and sensitive content filtering (Figure 2). To uniformly optimize language quality, entity coverage, and style consistency, a multi-task joint loss is introduced:

$$\mathcal{L}_{\text{joint}} = \mathcal{L}_{\text{gen}} + \alpha \mathcal{L}_{\text{ent}} + \beta \mathcal{L}_{\text{style}} \tag{5}$$

Among them, $\mathcal{L}_{\text{joint}}$ represents the overall loss of joint training, $\mathcal{L}_{\text{gen}}$ which is the loss generated by

ISSN 2616-7433 Vol. 7, Issue 8: 49-57, DOI: 10.25236/FSST.2025.070808

conditional language generation, which is related to $\mathcal{L}_{\text{ent}}$ the recall and accuracy of $\mathcal{L}_{\text{style}}$ cultural entities, and involves the classification or contrast loss of Lingnan style. Furthermore, $\alpha$ and $\beta$ are hyperparameters used to regulate the weights of entity and style subtasks.

In the online generation phase, a simple score weighting strategy is introduced to integrate the retrieval results and model confidence:

$$s_{final} = \gamma s_{ret} + (1-\gamma) s_{\text{lm}} \tag{6}$$

The final score used $s_{ret}$ here to rank candidate answers or candidate paragraphs $s_{final}$ is the relevance score obtained by the vector retrieval and re-ranking module, $s_{\text{lm}}$ which is the log probability or normalized confidence of the language model for the candidate output, $\gamma \in [0,1]$ and is the fusion coefficient of the retrieval signal and the generated signal.

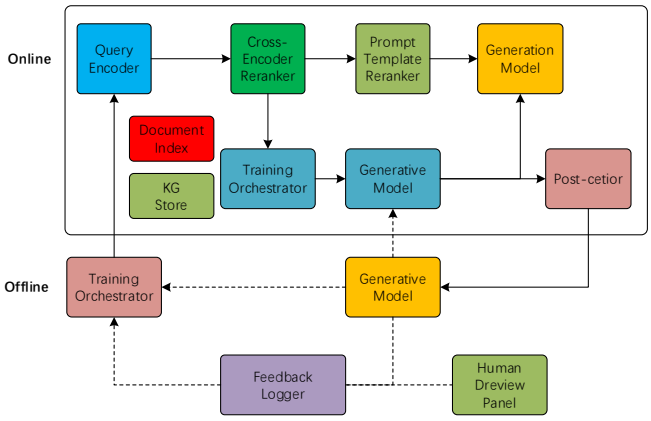


*Figure 2. Retrieval Enhancement Generation and Fine-tuning Architecture Driven by the Lingnan Architectural Culture Corpus*

## 4. Corpus Application and Experimental Evaluation

### *4.1 Experimental Design and Evaluation Task Definition*

Based on the aforementioned data and ontology, the evaluation framework covers three types of tasks: T1, Lingnan architectural culture Q&A, emphasizes factual accuracy and covers various cultural entities; T2, the generation of architectural and cultural descriptions focuses on coherence and stylistic unity; and T3, design assistance, emphasizes efficiency and satisfaction. The corpus employs stratified sampling of architectural examples to avoid single-region dominance in the test. T1 primarily consists of structured Q&A and explanatory paragraphs; T2 includes image/plan annotations and professional explanations; and T3 utilizes interaction logs collected through actual project workshops. The indicator system includes EM, F1, entity recall (T1), ROUGE, BERTScore, style confidence (T2), completion time, update prompts, and designer ratings (T3).

### *4.2 Experiment on Generating Knowledge of Lingnan Architectural Culture through Questions and Answers*

The knowledge-based question-answering experiment utilized structured corpora to form question-answer pairs and added content related to the evolution of ancestral halls, component functions, and ritual interpretations. Training used a unified retrieval-generation framework, contrasting with no retrieval, vector RAG, and overlay ontology weighting and knowledge graph embedding settings. During testing, the question distribution remained consistent, and generation length and temperature were standardized. Table 3 shows that as retrieval and ontology information were utilized, EM and F1 scores continuously improved, with the recall rate of cultural entities optimizing from below 50% to nearly 70%.

ISSN 2616-7433 Vol. 7, Issue 8: 49-57, DOI: 10.25236/FSST.2025.070808

*Table 3 Automatic evaluation results on Lingnan QA & explanation task*

| Model / Config | EM | F1 | BLEU | Cultural entity recall (%) |
|---|---|---|---|---|
| No Retrieval (baseline LLM) | 62.7 | 76.2 | 41.4 | 48.6 |
| BM25 + Generator | 68.9 | 80.3 | 46.7 | 53.1 |
| Dense RAG (bi-encoder) | 79.0 | 84.0 | 53.2 | 59.4 |
| RAG + Ontology Re-ranking | 81.2 | 85.3 | 56.8 | 66.4 |
| RAG + Ontology + KG Embedding | 82.5 | 86.1 | 57.6 | 68.9 |
| Distilled RAG + Ontology (small) | 80.4 | 84.7 | 55.9 | 64.2 |

As shown in Figure 3, when the corpus size is optimized from 0.5M tokens to 2.0M tokens, the RAG+ ontology model shows a steeper increase in both metrics compared to the baseline. This intuitively demonstrates the combined effect of the corpus and knowledge structure on the quality of question answering.

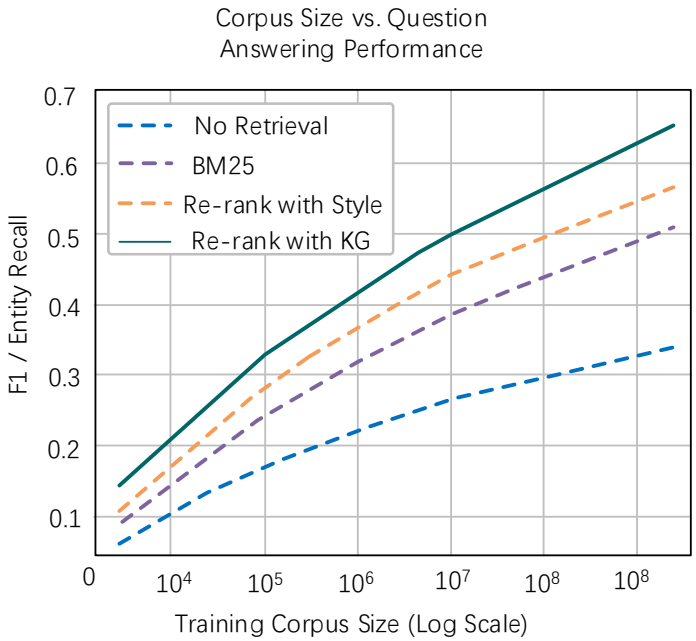


*Figure 3. Relationship between corpus size and question-answering performance*

### *4.3 Experiment on Generating Architectural Description and Cultural Interpretation Texts*

The architectural description generation experiment selected samples containing images, derived from survey reports and exhibition descriptions. Input was performed based on images and structured tags, with professional descriptions used as reference output content. When creating a general model, four types were set: template + terminology constraints, overlaid RAG, and RAG + LoRA fine-tuning. Two distillation models were also included for efficiency comparison. Results were evaluated using automatic metrics and blind review by three experts. Table 4 and Figure 4 show that after using domain-specific corpora and implementing style fine-tuning, the confidence scores for ROUGE-L, BERTScore, and Lingnan style were significantly improved, without excessive terminology stuffing.

*Table 4 Automatic metrics and style consistency scores on description generation*

| Model / Config | ROUGE-L | BERTScore | Lingnan style confidence (%) | Terminology density (%) |
|---|---|---|---|---|
| Generic Generator | 41.2 | 0.863 | 78.5 | 2.1 |
| + Prompt Template | 44.6 | 0.874 | 82.7 | 2.6 |
| + Domain Lexicon Constraints | 46.3 | 0.881 | 86.4 | 3.3 |
| RAG + Template | 48.9 | 0.892 | 89.8 | 3.1 |
| RAG + LoRA Style Tuning | 50.7 | 0.901 | 92.3 | 3.4 |
| Distilled RAG + LoRA (small) | 48.1 | 0.889 | 90.1 | 3.0 |

ISSN 2616-7433 Vol. 7, Issue 8: 49-57, DOI: 10.25236/FSST.2025.070808

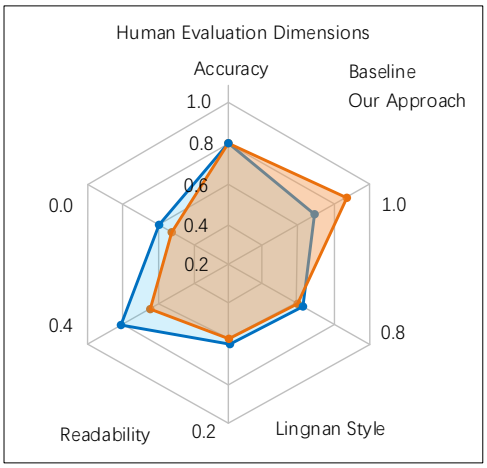


*Figure 4. Radar chart of human review dimensions (accuracy, readability, Lingnan style)*

### *4.4 Application Cases of Design Assistance and Form Generation*

The design-assisted experiment focused on three tasks: street renovation, arcade building restoration, and courtyard renewal. It compared three modes: plain text prompts, search-enhanced prompts, and a "template-guided + search" approach. Completion time, number of prompt updates, and text adoption rate were recorded, along with designers' evaluations of practicality and cultural conformity. Table 5shows that in complex tasks, the "template + search" approach significantly shortened update and completion times while maintaining high satisfaction and adoption rates.

*Table 5 Task completion and efficiency in design assistance scenario (template+RAG mode)*

| Task type | Avg completion time (min) | Avg prompt iterations | Designer satisfaction (1 – 5) | Adoption rate (%) |
|---|---|---|---|---|
| Streetscape Façade Guideline | 22.0 | 1.7 | 4.1 | 81 |
| Arcade restoration concept | 18.6 | 2.0 | 4.4 | 78 |
| Courtyard renovation scheme narrative | 26.5 | 2.5 | 4.5 | 84 |
| Heritage hotel lobby concept | 24.3 | 2.1 | 4.3 | 79 |
| Museum exhibition route description | 20.7 | 1.6 | 4.2 | 83 |
| Waterfront promenade design brief | 23.8 | 1.9 | 4.4 | 82 |

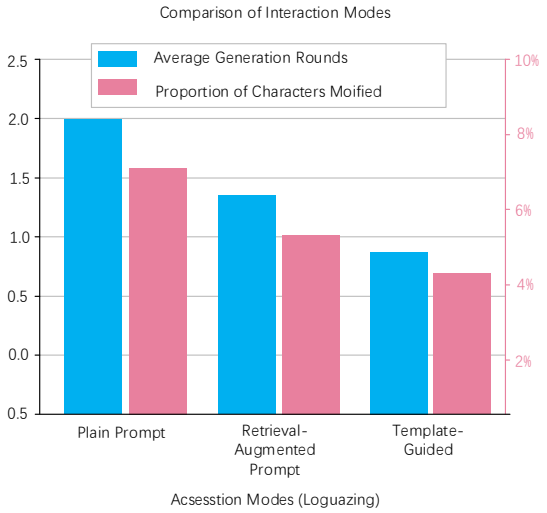


*Figure 5. Performance comparison of different interaction modes (plain text prompt / search-enhanced prompt / template-guided)*

ISSN 2616-7433 Vol. 7, Issue 8: 49-57, DOI: 10.25236/FSST.2025.070808

As can be seen from the comparison in Figure 5, enhanced retrieval and template guidance significantly reduced the proportion of manual modifications and the number of prompts, while maintaining or even slightly improving the adoption rate. This confirms that the corpus does have practical value in reducing cultural mismatch and optimizing the richness of inspiration.

### *4.5 Ablation Experiment and Error Analysis*

As shown in Table 6, to further quantify the contributions of different modules, the ablation experiment employed numerous settings based on the dimensions of knowledge structure and style constraints. Each setting underwent three repeated trials with automated metrics for tasks T1–T3, and the automated metrics were then averaged. During the ablation process, attention was paid to both the overall score and the recall rate of cultural entities and the accuracy of style classification, thereby identifying modules that were more important for "understanding Lingnan" and "resembling Lingnan."

*Table 6 Ablation results on three evaluation tasks*

| Config | T1–T3 joint score S | Cultural entity recall (%) | Style classification accuracy (%) |
|---|---|---|---|
| Full model (RAG + Ontology + KG + Style) | 82.3 | 68.9 | 91.7 |
| w/o Ontology Re-ranking | 77.6 | 60.4 | 90.1 |
| w/o KG Embedding | 78.9 | 62.8 | 88.3 |
| w/o Style Loss | 79.4 | 67.5 | 83.9 |
| 2/3 Lingnan subcorpus | 76.2 | 63.1 | 87.2 |
| 1/3 Lingnan subcorpus | 71.8 | 57.3 | 82.0 |

At the deployment level, it is necessary to balance the hardware environment and model size, so many inference schemes are created, including a complete cloud model, a local distillation model and a hybrid scheme, and end-to-end latency, traffic and resource consumption are recorded (see Table 7).

*Table 7 Inference efficiency and resource usage under different deployment schemes*

| Deployment scheme | Avg latency (ms) | Throughput (req/s) | GPU memory (GB) | CPU usage (%) |
|---|---|---|---|---|
| Cloud XL full model | 780 | 55 | 32 | 18 |
| Cloud L optimized | 540 | 72 | 20 | twenty four |
| Hybrid (cloud retriever, edge gen) | 620 | 68 | 16 | 31 |
| Edge distilled GPU | 410 | 81 | 8 | 37 |
| Edge distilled CPU-only | 1150 | twenty four | 0 | 76 |
| Shared GPU multi-tenant | 690 | 64 | 12 | 42 |

Figure 6 shows that knowledge structure reduction has a major impact on entity recall, while style loss reduction significantly reduces style accuracy. Different deployment schemes have a clear trade-off between efficiency and performance, which will provide a quantitative reference for subsequent applications in research areas and actual projects.

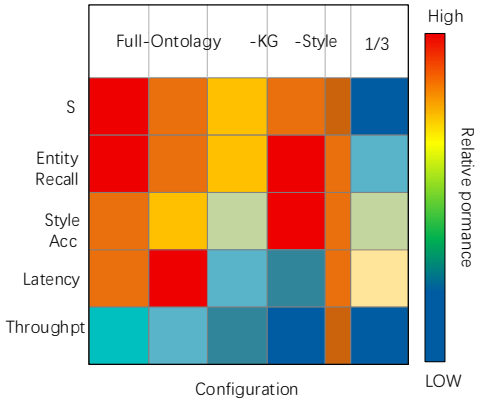


*Figure 6. Heat map of changes in various indicators under different ablation configurations*

ISSN 2616-7433 Vol. 7, Issue 8: 49-57, DOI: 10.25236/FSST.2025.070808

## 5. Conclusions and Outlook

This study established a Lingnan architectural culture corpus for generative AI, containing numerous source texts from areas such as Xiguan and the Ancestral Temple. It utilizes ontology and knowledge graphs to represent the connections between architecture, crafts, and cultural imagery, and integrates domain pre-training and graph embedding to achieve semantic unification. Experiments show that after adding the corpus and implementing retrieval enhancement, the accuracy of knowledge question answering and description generation, as well as the recall rate of cultural entities, are significantly improved. In design scenarios, RAG can reduce the number of prompt revisions and manual corrections. The limitation is that the existing corpus is primarily text-based, with limited 3D and interactive data. Future research will expand the multimodal corpus, strengthen cross-regional transfer and online updates to improve system robustness and generalizability.

## Acknowledgments

This work was supported by the 2023 Guangzhou Philosophy and Social Sciences Development "14th Five-Year Plan" Project (Project No. 2023GZGJ239), and by the 2024 Scientific Research Special Project of Guangdong Construction Polytechnic, titled "Construction and Application of a Lingnan Architectural Culture Corpus in the Context of Generative Artificial Intelligence" (Project No. KY2024-08).

2025 Teaching Reform Project of the Guangdong Provincial Teaching Steering Committee for Foreign Language Majors in Vocational Schools: "Pathways and Empirical Research on AIGC-Empowered Innovation in Foreign Language Teaching in Higher Education" (Project No. 2025wy06).

## References

*[1] Qing L, Wang H. A Study of the Construction Approaches to a Multimodal Narrative Discourse System for Lingnan Intangible Cultural Heritage[J]. Journal of Humanities, Arts and Social Science, 2024, 8(10).*

*[2] Zhang Y, Li W, Cai X. A cultural geography study of the spatial art and cultural features of the interior of Lingnan ancestral halls in the Ming and Qing dynasties[J]. Journal of Asian Architecture and Building Engineering, 2023, 22(6): 3128-3140.*

*[3] Qing¹ L, Wang H. Lingnan Cultural Image: A Case Study of Cantonese[C]//Proceedings of the 2025 3rd International Conference on Language, Innovative Education and Cultural Communication (CLEC 2025). Springer Nature, 2025, 938: 76.*

*[4] Zhong F. Sustainable transformation design of Lingnan vernacular architecture and landscape[M]//Frontiers in Civil and Hydraulic Engineering, Volume 1. CRC Press, 2023: 524-532.*

*[5] An H. Construction and Application of Corpora in Computer-Assisted Translation of Spanish Intangible Cultural Heritage for External Promotion[C]//2024 International Conference on Interactive Intelligent Systems and Techniques (IIST). IEEE, 2024: 490-494.*

*[6] Du D, Ding J, Liu Y. Knowledge graph construction of Chinese embroidery evolution based on associating cultural space and critical incidents under intangible cultural heritage[J]. The Electronic Library, 2025, 43(3): 283-302.*

*[7] Hou W, He Q, Li T, et al. The Method of Mining the Relationship Between the Use of Architectural Elements in Buildings and Cultural Connotation It Reflects: Case of Beijing's Representative Buildings[C]//International Conference on Human-Computer Interaction. Cham: Springer International Publishing, 2021: 74-87.*

*[8] Yazici M, Ozturk S D. An analysis of Rem Koolhaas's discourses on architecture and urban design using a corpus-based model[J]. Frontiers of Architectural Research, 2023, 12(2): 222-241.*

*[9] Elgibreen H, Faisal M, Al Sulaiman M, et al. An incremental approach to corpus design and construction: application to a large contemporary saudi corpus[J]. IEEE Access, 2021, 9: 88405-88428.*

*[10] Fan Z, Chen C. CuPe-KG: Cultural perspective–based knowledge graph construction of tourism resources via pretrained language models[J]. Information Processing & Management, 2024, 61(3): 103646.*