



Information Research – Vol. 31 No. 3 (2026)

Building a domain-specific large language model for Chinese ancient books: a case study on epitaph modelling

Yu Minjie, Jin Haoyue, Xue Shiliang

DOI: <https://doi.org/10.47989/ir31361795>

Abstract

Introduction. With the rapid advancement of generative artificial intelligence technology, large language models (LLMs) are expected to play an important role in improving access to and the use of library collections. This study constructs and develops a generative conversational large language model tailored specifically to epitaphs, overcoming traditional research paths and systems of digitisation.

Method. Based on the Chinese epitaph database built by the Zhejiang University Library and some commercial epitaph databases, an epitaph generative dialogue large language model was constructed using Doubao as the base model through pre-training, supervised fine-tuning, and other full process trainings.

Analysis. This study used human evaluation, in which experts, researchers, or general users assessed the quality of model-generated outputs. The evaluation included the epitaphs of Cui Yi, Liu Yi, and Lihua in the Tang dynasty.

Results. The model successfully performed intelligent interpretation of epitaphs, automatic information extraction, and other knowledge-based services, thereby providing new ideas for constructing an ancient literature corpus from excavated materials represented by epitaphs.

Conclusion. A large-scale domain-specific foundational model tailored for library applications was successfully developed. The study offers innovative approaches for processing historical texts and suggests new directions for AI integration in library knowledge systems.

Introduction

Ancient Chinese books and documents contain essential records of China's history and cultural heritage. They integrate the cultures and thoughts of numerous ethnic groups to form philosophical systems, such as Confucianism, Taoism, and Legalism. These texts, refined over centuries, have contributed to the richness of traditional Chinese culture.

In April 2022, the General Office of the Central Committee of the Communist Party of China and the General Office of the State Council issued *Opinions on Advancing Ancient Books Work in the New Era*, outlining objectives such as 'strengthening interdisciplinary integration, incorporating fields such as materials technology, information technology, and artificial intelligence' and 'enhancing the circulation and collaborative management of ancient books data to facilitate the aggregation and sharing of digitised ancient books resources' (Central Committee of the Communist Party of China & General Office of the State Council, 2022, paras. 4–5).

In the current field of ancient book digitisation, most operations rely on scanned transcriptions of ancient texts, including segmentation, punctuation, and title recognition, which are then structured into searchable databases. However, the extraction of key information such as personal details, locations, event attributes, and temporal data from these texts remains unexplored, limiting users' ability to effectively access and use the fragmented knowledge contained therein.

With advancements in large-language modelling technology, the boundaries of natural language processing tasks have expanded, enabling diverse downstream tasks to be accomplished using a unified prompting paradigm. General-purpose large language models (LLMs) such as [ChatGPT](#), [Qianwen](#), [Zhipu Qingyan](#), and [Doubao](#) exhibit excellent performance in understanding and generating modern Chinese. However, their ability to comprehend Classical Chinese remains uncertain. Through extensive evaluation of state-of-the-art large language models, Zhao et al. (2025) revealed significant performance gaps between understanding and generation tasks, with models achieving promising results in comprehension but struggling considerably with generation tasks, particularly those requiring deep cultural knowledge and adherence to classical formats. Ancient texts, as a vertical domain with a high reading comprehension threshold, pose unique challenges owing to the peculiarities of Classical Chinese and textual characteristics, such as variant characters, mixed use of simplified and traditional characters, and complex chapter structures. The applicability of general large language models to ancient texts requires further exploration. Moreover, these models operate on large datasets, demand high technical expertise, exhibit strong adaptability, and incur significant deployment and operational costs.

This study develops a generative conversational large language model tailored to epitaphs, intending to move beyond traditional research paths and systems of digitisation. The objective is to transform the specialised database services from data-oriented to knowledge-oriented approaches while expanding their accessibility. By intelligently uncovering insights into family heritage, marital alliances, life experiences, political activities, literary endeavours, and the social interactions of epitaph subjects, this study enhances knowledge-based and intelligent services. Additionally, it provides a reference framework for constructing domain-specific large models based on ancient library book resources.

Literature review

Libraries and epitaphs: traditional practices before the large language model era

Before the popularisation of the large language model, libraries had long been responsible for the core work of collecting, organising, digitising, and providing public services for ancient inscriptions and documents such as epitaphs. The traditional method of handling epitaph documents in libraries involves manual sorting, preservation of rubbings, bibliographic cataloguing, and standardised metadata construction.

Library professionals and researchers rely on traditional theories of epigraphy and philology to finely organise scattered epitaph rubbings and historical documents, compile epitaph catalogues and specialised literature collections, and complete basic proofreading and historical verification work. Peking University Library, for example, compiled the *Catalogue of epitaph rubbings from various dynasties* based on its collection of epitaphs (Hu & Tang, 2013). In the early stage of its digital transformation, the library digitised its epitaph resources through scanning, optical character recognition, database construction, and other methods, enabling online retrieval and static resource display of epitaph literature. Based on this collection of epitaphs, Zhejiang University Library digitised them and built a database of epitaphs from various dynasties in China. However, the traditional library processing mode has inherent shortcomings and bottlenecks: manual sorting and text annotation are time consuming and labour intensive and rely on the professional accumulation and practical experience of senior researchers. Furthermore, traditional digital systems can only achieve basic resource storage and keyword matching retrieval; they do not have advanced processing capabilities such as deep semantic analysis, intelligent information extraction, automatic translation, and text summarisation for ancient stele texts (Li, 2019). The long-standing technological deficiencies and efficiency challenges have constrained the deep development and intelligent use of epitaph literature resources. Therefore, there is an urgent need to rely on artificial intelligence innovation technology to empower the inheritance and academic research of the cultural heritage of epigraphy.

Current practices of large language models in vertical domains

Owing to differences in the external knowledge characteristics between general and vertical domains, directly applying publicly available, knowledge-injected pre-trained language models to vertical domains often produces suboptimal results. Academia and industry are actively exploring strategies to adapt these models to specific industries and application scenarios, enabling the customisation of large models for vertical domains. Based on the degree of customisation, from low to high, or conversely, the extent of reliance on general large models, from high to low, these adaptations can be categorised into three types: (1) vertical-domain enhancement using the full architecture of general large models, (2) vertical-domain large model transformation leveraging pre-training modules, and (3) full architecture construction of vertical-domain large models without pre-training modules. Representative vertical-domain-specific models, such as AlpaCare (Zhang et al., 2023), HuatuoGPT (Zhang et al., 2023), Lawyer LLaMA (Huang et al., 2023), and OpenBioMed (Luo et al., 2024), have emerged in fields such as law, medicine, and agriculture. A consistent finding across these studies is that simply fine-tuning general models with domain data is often insufficient; deeper architectural or training adaptations are required to achieve expert-level performance.

Research on the application of large language models in libraries is extensive and primarily focuses on insights, suggestions, and strategies for library transformation and innovation to support future development. This includes discussions on leveraging ChatGPT to improve reference services, cataloguing, and reading promotion, among other business scenarios. However, as is detailed later, these library-oriented studies rarely address the specific challenges of handling ancient Chinese epitaphs, nor do they explore how libraries traditionally managed such materials before the advent of large language models.

Recent evaluations of Chinese general-purpose large language models, such as Qianwen and Baichuan, have shown that instruction tuning is critical for unlocking their potential in complex tasks (Zhao, et al., 2024). Nevertheless, when applied to ancient Chinese texts, these models often face significant limitations due to domain-specific vocabulary, syntactic structures, and a lack of contextual understanding. This requires the development of vertical-domain models that can integrate domain knowledge more deeply, addressing the limitations of a one-size-fits-all approach.

In traditional Chinese cultural research, technologies such as automatic punctuation, annotation, translation, and collation of ancient texts have garnered considerable attention. However, the necessary infrastructure for these research efforts, such as large corpora and corresponding knowledge auxiliary tools, remains insufficient. Chen (2022) proposed that to improve the character recognition rate of computers for ancient texts, it is essential to integrate various technologies, including corpora, word lists, knowledge bases, and artificial intelligence supplemented by manual proofreading. Among these, corpora and word-list knowledge bases play crucial roles, serving as libraries for other aspects of traditional cultural research. Against this backdrop, domain-specific large vertical models have been developed for ancient texts. The deep integration of large language models with intelligent ancient text processing holds substantial cultural value and is capable of significantly improving the efficiency of organising and researching ancient texts while promoting the intelligent processing of ancient text information. Table 1 presents a selection of large language models derived from ancient Chinese books, illustrating a growing trend of domain specialisation in this area.

Large language model	Development unit	Main function	Application scenario
Xunzi series	Nanjing Agricultural University and Zhonghua Book Company	Ancient text understanding, generation, translation, and annotation	Intelligent indexing of ancient books, information extraction, poetry creation, translation and understanding
TongGu	South China University of Technology	Translation of literary and vernacular Chinese, punctuation, and ancient book search	Digitalisation, intelligent development and utilisation of ancient books
Hanwang Ancient Chinese	Hanwang Technology	Deep learning capabilities, ancient Chinese text recognition, comprehension and generation	Digitisation of ancient books, arrangement and research of ancient books, construction of smart libraries
Huawei Shennong Brain	Huawei Technologies	Integrate data in the field of traditional Chinese medicine based on artificial intelligence technology	Intelligent diagnosis and decision support, research and development of traditional Chinese medicine, and optimal allocation of medical resources
Digital Qihuang	East China Normal University, Shanghai University of Traditional Chinese Medicine	Combining Chinese and Western medicine, comprehensively integrating Chinese and Western medicine data and knowledge	Intelligent Q&A and consultation, assisting medical decision-making, and promoting innovation in traditional Chinese medicine
AI Taiyan	Beijing Normal University	Word interpretation, literary and colloquial translation, sentence reading and punctuation, and allusion analysis	Ancient book compilation, teaching and scientific research, language tool development

Table 1. Large language models derived from ancient Chinese books

Several research groups have demonstrated the feasibility of building vertical large language models for ancient texts. Zuo et al. (2024), leveraging the [Xunzi](#) ancient text large language model, used unannotated corpora from the historical and classical sections of the *Siku Quanshu* as input. They approached tasks such as lexical analysis, entity recognition, relation extraction, text classification and matching, and text translation in ways closely aligned with the original processing of ancient texts and contextual understanding similar to that of human experts to achieve automatic processing of ancient texts.

Comparative studies have demonstrated that in vertical domains, large language models trained professionally perform better and hold more value in specific domain applications than general-purpose base models. Research shows that the Xunzi model achieves substantial improvements

over general models on domain-specific benchmarks (Zuo et al., 2024). Zhang et al., building upon the open-source LLaMA model, incorporated corpora from traditional Chinese medicine textbooks and various traditional Chinese medicine websites. Through continued pre-training, supervised fine-tuning, and direct preference optimisation, they developed [Huang-Di](#), a generative conversational large language model tailored for traditional Chinese medicine ancient texts. Automatic and manual evaluations confirmed its superior performance in ancient traditional Chinese medicine text domains. Unlike traditional knowledge-based ancient text search engines, Huang-Di does not rely on keywords or phrases (Zhang et al., 2024). Users can engage in one-on-one interactions with the model using an easily understandable modern language instead of Classical Chinese. This significantly lowers barriers to accessing specialised knowledge while meeting various knowledge needs, including ancient text interpretations, traditional Chinese medicine consultations, and health maintenance guidance.

This shift from keyword-based retrieval to conversational interaction represents an important evolution in making domain-specific knowledge more accessible (Zhang et al., 2024).

Li et al. (2024) developed [AI Taiyan](#), a Classical Chinese large language model, providing a foundational reference for building specialised large language models from scratch. The development process began with designing a small-scale large language model architecture based on domain knowledge learning requirements and available data. The AI Taiyan model demonstrates strong capabilities in interpreting classical literature, supporting various challenging Classical Chinese understanding tasks such as sentence punctuation, allusion identification, word explanation, and translation between Classical and vernacular Chinese. It supports both simplified and traditional Chinese texts and holds substantial potential for application in ancient text organisation, dictionary compilation, revision, and language research.

Wu et al. (2024), using biographical materials from the *Twenty-Four Histories* as case studies, explored feasible approaches to applying automatic summarisation to ancient texts using artificial intelligence generated content technology and both extractive and generative methods. They reported that although general large language models exhibit unique capabilities in automatic summarisation for ancient texts, they struggle to distil key ideas effectively. This finding is particularly relevant for epitaphs, which are highly structured and information-dense, suggesting that domain-adapted summarisation mechanisms are needed.

Technical barriers in large-scale models for ancient texts

Currently, the corpora used to train generative AI models for ancient texts are undergoing evolution and expansion. Domain-specific large models such as Xunzi (Wang, 2023) and AI Taiyan (Li et al., 2024) show clear advantages over general-purpose models such as Qianwen, ChatGPT, and [DeepSeek](#) in tasks including Classical Chinese translation and automatic punctuation. However, these models remain limited in their ability to deeply interpret the connotations of ancient vocabulary and identify causal relationships between entities. These limitations are due to the over-representation of transmitted texts (for example, Confucian classics, historical records, and philosophical writings) and the under-representation of excavated materials, such as epitaphs, in their training data. The resulting gaps occasionally lead to inaccuracies in output. As Song et al. (2025) noted in their comprehensive survey on knowledge injection, integrating domain-specific knowledge into large language models is a critical task that directly addresses such under-representation.

As a critical category of historical documents, epitaphs function not only as material carriers of literary texts but also as direct objects of literary research. They contain abundant biographical, genealogical, and geographical information often absent from official histories, thereby offering substantial historical and cultural value. Although high-quality training data are considered essential, the total number of digitised and optical character recognition-processed epitaphs

remains relatively small. The development of a specialised variant font library in conjunction with epitaph studies, alongside collaboration between domain experts and librarians, could facilitate the construction of a high-quality corpus to address this gap.

With the rapid advancement of generative large language models, the field of ancient Chinese texts has witnessed the emergence of domain-specific models in areas such as traditional Chinese medicine and the collation of ancient books. However, the development of large models dedicated specifically to epitaphs remains unexplored. As essential historical materials, epitaphs function both as physical carriers of literary works and as primary sources for historical and literary research, offering a rich repository of historical and cultural information.

Thus far, no studies have addressed the creation of large-scale language models tailored to epitaphs. Constructing such a model would enable deeper mining and analysis of ancient textual resources in this domain, reveal underlying linguistic and historical patterns, and support the transformation of ancient texts into structured knowledge. This effort would also promote the reactivation and reinterpretation of their cultural value by extracting and reorganising knowledge units tied to traditional culture, ultimately fostering the innovative development of China's outstanding cultural heritage.

Constructing a large language model for epitaphs

Characteristics of epitaph literature

Epitaphs are a literary form used to commemorate the deceased, typically consisting of two parts: the epitaph proper (*zhìwén*) and the inscription (*míngwén*). As unique genres, epitaphs exhibit characteristics in content, structure, stylistic features, medium, form, and historical evolution.

Diverse literary forms

While maintaining overall stability, the epitaphs display rich diversity. Styles vary across different eras and social strata, reflecting the profound impact of temporal changes and class differences on literary forms. During the Tang and Song dynasties, epitaphs evolved from parallel prose to prose, marking a significant shift from mourning literature to biographical literature, enhancing their literary nature and realism.

During the Song dynasty, epitaphs became vital materials for historiography. In terms of content, their use expanded beyond the nobility to the common people. This trend continued and intensified in later periods, particularly during the Ming dynasty, leading to the proliferation of epitaph writing. Unlike noble epitaphs, commoner epitaphs focused more on daily life and the inner world, in contrast to earlier records that emphasised official positions and achievements. Understanding this diversity will aid the development of large language models that can provide customised services tailored to various subject areas.

A relatively fixed structure

From the Wei, Jin, Southern, and Northern dynasties to the Sui and Tang dynasties, epitaphs developed standardised structures and established writing conventions. By the time of the Southern dynasties' *Anthology of Literature* (*Wénxǔǎn*), epitaphs were formally recognised as an independent literary genre (Xiao, ca. 520–530 C.E./2025). Viewed synchronically, the content of epitaphs aligns with the thirteen items proposed by the Ming scholar Wang Xing (n.d.): posthumous names (*shìhào*, names that reflected official assessments of deceased individuals' political careers), courtesy names (*zì*, alternative formal names bestowed upon elite Chinese men upon reaching adulthood and, occasionally, upon women), surnames, native places, family origins, official positions, life experiences, dates of death, ages, spouses, children, burial dates, and burial locations. Epitaphs maintained a consistent framework comprising the epitaph proper and the inscription (Wang, n.d.). Diachronically, from their emergence in the Wei and Jin periods to their decline in the late Qing and early Republican eras, epitaphs remained highly stable. Early epitaphs often

employed parallel prose, emphasising formal beauty. From the mid-Tang period onward, however, they transitioned toward prose, marking a shift from formalism to content-driven expression. Despite these stylistic transformations, their fundamental structures and characteristics remained largely unchanged, distinguishing them from other literary forms.

Abundant existing corpora

Rough estimates suggest that more than 15,000 epitaphs from the Northern and Tang dynasties alone have been documented, with numerous new discoveries each year owing to ongoing construction and archaeological excavations. Major collections such as the *Complete Prose of the Tang* (*Quán Táng Wén*) and various anthologies contain many preserved epitaphs, with only a few overlapping with recently excavated ones. Two source types are examined: transmitted versions, which refer to epitaph texts preserved in later compiled historical works and literary collections, and stone rubbings, which are impressions taken directly from original stone inscriptions. Comparisons between these two types revealed minimal textual differences, indicating the high textual reliability of transmitted epitaph records.

Moreover, the number of epitaphs from the Song, Yuan, Ming, and Qing dynasties far surpasses those from the Tang dynasty, although scholarly research on later era epitaphs remains in its early stages. This vast corpus provides a solid foundation for the development of large language models, enabling in-depth linguistic and historical analysis while supporting the structured digitisation of ancient textual resources.

Clear information markers

Beyond their literary functions, epitaphs serve as valuable materials that supplement and corroborate historical records. Because of their unique genre and structured composition, the information they contain is often explicit and systematically organised, as reflected in Wang Xing's thirteen items. This standardised structure facilitates the systematic extraction, integration, and categorisation of epitaphic data into relevant historical databases.

The formulaic nature of epitaphs significantly enhances information retrieval, as different types of information are typically indicated by specific markers. For example, official titles are often introduced by terms such as 'appointed' (*chú*), 'promoted' (*qiān*), 'transferred' (*zhuǎn*), or 'granted' (*zèng*); geographical information by 'died' (*zú*), 'buried' (*zàng*), 'interred' (*biǎn*), or 'accompanied in burial' (*fù*); and temporal information by 'at the age of' (*xiǎngnián*), 'lived to' (*xiǎngshòu*), or 'spring and autumn' (*chūnqiū*).

Using large language models to extract and analyse information from epitaphs can provide researchers with key data on the deceased and their families, helping to elucidate relationships and social connections, thereby presenting a clear and intuitive knowledge system.

Selection of base models

Doubao, Zhipu Qingyan, DeepSeek, and Qianwen are the four major models that support Chinese language training. During data extraction stage, the commercial application programming interface of these models was used. To select the most suitable base model for domain adaptation to epitaphs, the models' performance was evaluated under zero-shot conditions, that is, without any domain-specific fine-tuning.

Evaluation dataset. We selected three epitaphs from different dynasties: one from the Tang dynasty, one from the Northern Song dynasty, and one from the Ming dynasty. These epitaphs were chosen to assess the ability of the model to generalise across historical periods and text variations.

Evaluation tasks and scoring criteria. Each model was evaluated on four tasks: sentence segmentation, logical judgment, translation, and summary. All outputs were scored on a ten-point

scale (1 = completely incorrect, 10 = fully accurate). The scoring criteria for each task is presented in Table 2.

Task	Score	Criteria
Sentence segmentation	10	all punctuation correct
	8–9	1–2 errors
	6–7	3–5 errors
	≤5	more than 5 errors
Logical judgment	10	all logical relationships correctly identified
	8–9	minor errors
	6–7	2–3 errors
	≤5	major errors
Translation	10	fully accurate and fluent
	8–9	1–2 minor inaccuracies
	6–7	3–5 errors
	≤5	significant errors
Summary	10	all key information captured
	8–9	missing minor details
	6–7	missing 2–3 key details
	≤5	missing most key information

Table 2. Evaluation tasks and scoring criteria for models

Evaluation Procedure. Three domain experts with PhDs in Chinese philology independently evaluated the outputs. The final score for each model–task–epitaph was the average of the three experts' scores. Inter-rater reliability was high (intraclass correlation coefficient = 0.87).

Results. Tables 3–6 present the scores of the four models. Except for logical judgment ability, where Doubao and DeepSeek scored equally (8.7), Doubao outperformed the other models in the remaining three aspects. Additionally, Doubao had the smallest variance in scores across all four dimensions, indicating more consistent performance in interpreting epitaphs from different dynasties and demonstrating better generalisation. Based on the overall scores across all four aspects, Doubao was selected as the foundation for our model.

Large language model	Epitaph (Tang dynasty)	Epitaph (Northern Song dynasty)	Epitaph (Ming dynasty)	Mean score	Variance
Doubao	9	9	9	9	0
Zhipu Qingyan	9	9	8	8.7	0.2
DeepSeek	9	9	8	8.7	0.2
Qianwen	9	9	8	8.7	0.2

Table 3. Sentence breaking ability of large language models (Full Score: 10 points)

Large language model	Epitaph (Tang dynasty)	Epitaph (Northern Song dynasty)	Epitaph (Ming dynasty)	Mean score	Variance
Doubao	9	9	8	8.7	0.2
Zhipu Qingyan	5	5	9	6.3	3.6
DeepSeek	9	9	8	8.7	0.2
Qianwen	5	9	9	7.7	3.6

Table 4. Logical judgment ability of large language models (Full Score: 10 points)

Large language model	Epitaph (Tang dynasty)	Epitaph (Northern Song dynasty)	Epitaph (Ming dynasty)	Mean score	Variance
Doubao	9	8	9	8.7	0.2
Zhipu Qingyan	6	8	9	7.7	1.6
DeepSeek	7	9	8	8	0.7
Qianwen	7	8	7	7.3	0.2

Table 5. Translation ability of large language models (Full Score: 10 points)

Large language model	Epitaph (Tang dynasty)	Epitaph (Northern Song dynasty)	Epitaph (Ming dynasty)	Mean score	Variance
Doubao	9	9	9	9	0
Zhipu Qingyan	8	9	9	8.7	0.2
DeepSeek	9	9	8	8.7	0.2
Qianwen	9	6	4	6.3	4.2

Table 6. Summary and induction ability of large language models (Full Score: 10 points).

Model construction

Data collection and reprocessing

This study primarily collected data from the [Database of Chinese Epitaphs through the Ages](#) developed by the Zhejiang University Library's Center for Ancient Books and Stone Rubbings Research and Preservation, as well as from certain commercial epitaph databases. The data consisted predominantly of epitaphs from the Northern, Tang, and Five Dynasties.

Dataset construction

To construct a high-quality dataset for domain-specific model development, this study organised the collected data into two complementary components:

1. **Epitaph corpus (domain knowledge source):** A collection of 6,800 epitaph texts and rubbings. Each text underwent cleaning, classification, and annotation to ensure data quality and consistency. The annotated metadata included tomb owner name, burial date, official titles, and other relevant biographical information. This corpus served as the foundation for continued pre-training and knowledge base construction.
2. **Instruction tuning dataset (task-specific data):** A set of 20,000 instruction-response pairs was derived from the epitaph corpus. Specifically, each epitaph was used to generate multiple instruction-response pairs covering tasks such as sentence segmentation, translation, information extraction, and summarisation. These pairs were constructed through three methods: template-based generation (approximately 12,000 pairs), self-instruct augmentation (approximately 6,000 pairs), and manual Q&A pair creation (approximately 2,000 pairs).

The relationship between the two components is as follows: the 6,800-entry epitaph corpus provides the domain knowledge, while the 20,000 instruction-response pairs adapt this knowledge to specific task formats for supervised fine-tuning.

Data structuring and annotation

To enhance the model's ability to comprehend and process epitaph texts, this study first established a knowledge structure based on the characteristics of epitaphs and research requirements, designing a reasonable framework that classifies the data by dynasty, region, and social status of individuals. For the epitaph texts, seven fields were defined and provided professional descriptions to support model understanding:

Title	Headings located at the beginning of the epitaph
Inscriber	The person who inscribed the epitaph, often indicated by terms such as <i>shu</i> (书, scribe) or <i>shudan</i> (书丹, inscription)
Author	The author of the epitaph text, commonly denoted by words such as <i>zhuan</i> (撰, compose), <i>wen</i> (文, text), <i>ci</i> (词, words), <i>zhuanwen</i> (撰文, compose text), or <i>shu</i> (述, narrate)
Dynasty	The dynasty during which the tomb owner was buried
Chinese era date	The traditional Chinese calendar date (year, month, and day) of the tomb owner's burial is typically expressed using era names
Gregorian date	The Gregorian calendar date corresponding to the Chinese era date is numerically represented
Full text	The complete text of the epitaph, including title, main body, and colophon

Table 7. Epitaph text fields

The data were structured in this manner to enhance the model's ability to comprehend and process epitaph texts and improve its performance in tasks related to the analysis and interpretation of ancient Chinese inscriptions.

Dataset characteristics

To characterise the epitaph corpus and demonstrate the complexity of the domain, a statistical analysis of the 6,800 texts was conducted. Table 8 summarises the key characteristics.

Characteristic	Value
Total number of epitaphs	6,800
Total characters	Approx. 4.2 million
Average length per epitaph	618 characters (SD: 342)
Shortest epitaph	48 characters
Longest epitaph	2,156 characters
Vocabulary size (unique words)	Approx. 18,500
Average sentence length	12.4 characters
Domain-specific terms	Approx. 1,200 (official titles, toponyms, funeral formulae)

Note. SD = standard deviation.

Table 8. Statistical characteristics of the epitaph corpus

The length distribution reveals that the majority of epitaphs (62%) fall within the range of 300–800 characters, while short epitaphs (<300 characters) and long epitaphs (>800 characters) account for 18% and 20%, respectively. This diversity in length ensures that models trained on this corpus can handle a wide range of text complexities. The vocabulary comprises approximately 18,500 unique words, with a substantial proportion (approximately 1,200 terms) being domain-specific terminology related to official titles, toponyms, and funeral formulae. This high density of specialised terms underscores the necessity of domain-specific modelling for accurate processing.

Sample training

To meet the technical requirements for developing domain-specific models and to ensure research reproducibility, the training process in this study was conducted using the enterprise-grade

machine learning framework provided by Volcengine. This platform offers hosted infrastructure for distributed training and large-scale model deployment. Figure 1 illustrates the key technologies and implementation pathways in the process of constructing the epitaph model, from base model selection to full development. Building on the Doubao general-purpose foundational model, this study implemented three critical strategies: continued pre-training, instruction tuning, and retrieval-augmented generation supported by a local knowledge base.

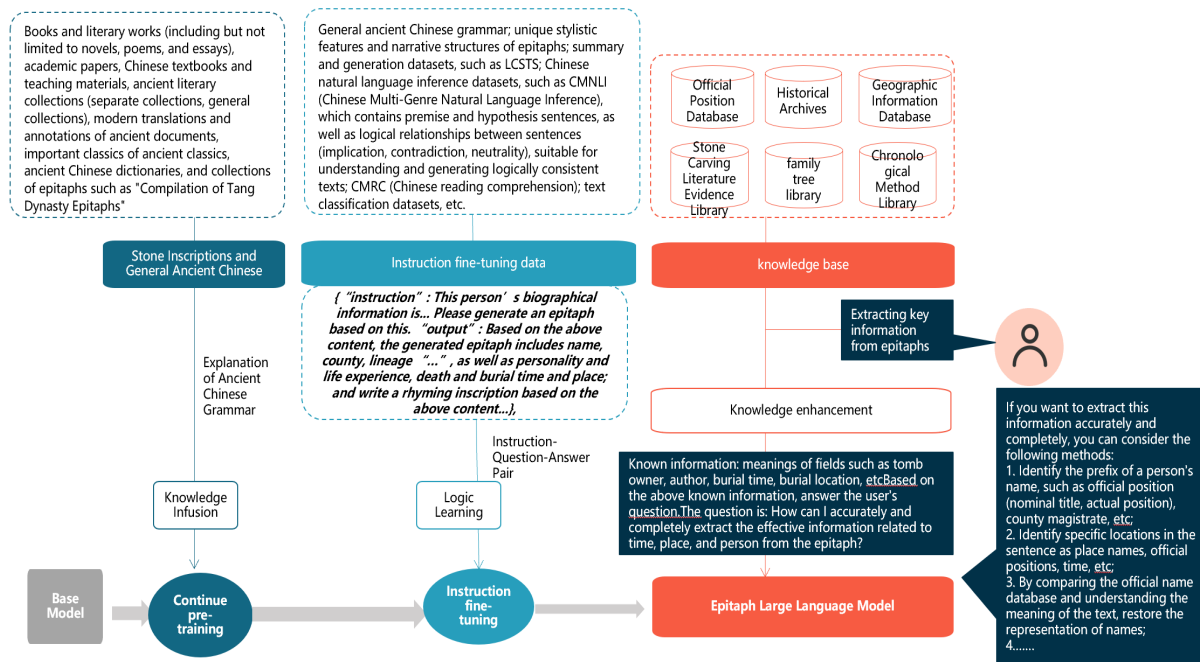


Figure 1. Key technologies and implementation approaches for constructing epitaph models.

To further enhance the model's proficiency in the epitaph domain, this study implemented two key strategies: continued pre-training and instruction fine-tuning. Complementing these training strategies, retrieval-augmented generation uses a local knowledge base as an inference-level enhancement module.

1. **Continued pre-training on epitaph corpus.** The goal of continued pre-training is to inject domain-specific knowledge into the general model, enabling a better understanding of the linguistic features and patterns specific to epitaphs.
 - a. **Data preparation:** To prevent catastrophic forgetting while preserving the general ability of the underlying model, this study mixed general domain data such as ancient literary texts with a specialised corpus of epitaphs during the continued pre-training phase. The cleaned and organised epitaph data was converted into standard JSON lines format, which is {"text": "xxxxx"}.
 - b. **Training configuration:** Using the pre-training tools provided by the Volcengine platform, the study employed a parameter-efficient fine-tuning method, particularly low-rank adaptation (LoRA). The model was trained on an NVIDIA A100 GPU with a learning rate of 2e-5 and a batch size of 16, for a total of 3 epochs. This balanced approach enabled the model to effectively use high-quality domain data, thereby enhancing its understanding and processing capabilities for tasks related to epitaphs.
2. **Instruction fine-tuning based on epitaphs.** Instruction fine-tuning involves further training the large language model on a dataset composed of instruction-response pairs in a supervised manner. This process bridges the gap between the model's next-word

prediction objective from pre-training and the goal of following human instructions, thereby improving its comprehension and execution capabilities in handling epitaph-related tasks.

- a. **Dataset construction:** The core of instruction fine-tuning is to create a high-quality, large-scale, and domain-specific dataset. The data were further processed and converted into a conversational JSON format: {"messages": [{"role": "user", "content": "Question"}, {"role": "assistant", "content": "Answer"}]}. Three key methods were used to construct this dataset for epitaphs:
 - i. **Template-based data integration:** Using pre-defined templates to systematically generate instruction-response pairs from the epitaph corpus.
 - ii. **Self-instructed data augmentation:** Using the model itself to generate additional instruction-response pairs, thereby expanding the dataset and introducing diversity.
 - iii. **Document-based Q&A generation:** Asking questions based on the content of the epitaph to generate corresponding answers and create an instruction-response dataset.
 - b. **Training configuration:** With the supervised fine-tuning function of the Volcengine platform, this study deepened the model's understanding of the field of epitaphs. The model was optimised using the AdamW optimiser, setting the rank of LoRA to 8 and the learning rate to $1e-5$, training it for 5 epochs with a warm-up ratio of 0.05.
3. **Retrieval-augmented generation and local knowledge base.** Finally, by integrating local knowledge bases and enhancing retrieval generation techniques, this study improved the vertical domain performance of the model:
- a. **Knowledge base construction.** The structured metadata of epitaphs (such as official records, geographic data, and historical chronology) was divided into text blocks of approximately 400 tokens. The Doubao Embedding model was used to vectorise these text blocks and establish an index in the vector database.
 - b. **Retrieval mechanism.** During the inference stage, the user's query was transformed into a vector, and the top two most relevant historical records were retrieved through cosine similarity search. By dynamically concatenating the retrieved context with prompt words, this significantly improved the model's ability to extract key information such as the tomb owner's name, life story, and time of writing, and successfully associated this information with validated historical knowledge, thereby preventing AI from producing hallucinations.

Evaluation

To evaluate the performance of our domain-specific epitaph model, a human evaluation was conducted.

Evaluators

Three domain experts participated in the evaluation. All evaluators held doctoral degrees in Chinese Philology or Classical Chinese Literature, with specialised expertise in Tang dynasty epitaphs and palaeography. Each evaluator had more than five years of experience in the transcription, annotation, and interpretation of ancient Chinese epitaph texts. Prior to the formal evaluation, each annotator completed a training session comprising twenty example to ensure a consistent understanding of the evaluation criteria.

Baseline models

To assess the necessity of domain-specific fine-tuning and to contextualise the performance of the proposed model, it was compared against three representative general-purpose large language models:

1. **Doubao (base).** The foundational model upon which our epitaph model is fine-tuned. This comparison directly isolates the impact of domain-specific instruction tuning.
2. **Qianwen.** A state-of-the-art Chinese general-purpose large language model developed by Alibaba, widely used for Chinese language tasks.
3. **ChatGPT (GPT-4).** An internationally prominent general-purpose large language model, included to evaluate cross-domain and cross-linguistic performance on Classical Chinese texts.

All baseline models were evaluated under zero-shot conditions, without domain-specific fine-tuning or task-specific training.

Evaluation tasks and scoring criteria

The evaluation assessed the model's performance across four task types, each reflecting a core capability required for epitaph processing. For each task, evaluators rated the model's output on a ten-point scale (1 = completely incorrect or unusable, 10 = fully accurate with no errors). The four tasks were defined as follows:

1. **Sentence segmentation.** The ability to correctly segment Classical Chinese text into meaningful clauses and sentences, adhering to syntactic and semantic conventions of Tang dynasty epitaphs. Scores were based on the accuracy of segmentation boundaries and the preservation of original textual meaning.
2. **Logical judgment.** The ability to identify and reason about logical relationships within epitaph texts, including temporal sequences (for example, burial dates), kinship relations, and official career progressions. Scores reflected the accuracy of extracted relational information.
3. **Translation.** The ability to translate Classical Chinese epitaph text into modern Chinese while maintaining faithfulness to the original meaning, preserving domain-specific terminology (for example, official titles, funeral formulae), and ensuring fluency in modern Chinese expression.
4. **Summary and induction.** The ability to generate concise summaries that capture key information from epitaph texts, including the deceased's identity, family background, life events, and burial details, without omitting critical information or introducing hallucinations.

Evaluation materials

Three Tang dynasty epitaphs, those of Cui Yi, Liu Yi, and Li Hua, were selected as evaluation materials. These epitaphs were chosen based on the following criteria: (1) they were representative of typical Tang dynasty epitaph structure and content; (2) they vary in length (short, medium, and long) to assess model robustness across text complexity; and (3) their authoritative transcriptions and interpretations have been validated by professional researchers in the field, ensuring reliable ground truth for evaluation.

Evaluation procedure

Each evaluator independently assessed the model's outputs for all three epitaphs across the four task dimensions. The final score for each epitaph-task combination was determined by the evaluator based on the ten-point rubric. To ensure consistency, evaluators were provided with a detailed scoring guideline with annotated examples for each score level. Inter-rater reliability was assessed using the intraclass correlation coefficient, yielding a value of 0.89 (95% CI: 0.82–0.94),

indicating excellent agreement among evaluators. Discrepancies in scoring were resolved through consensus discussion. Evaluators were blinded to model identity to mitigate potential evaluator bias.

Results

Table 9 presents the evaluation results for our epitaph model. Table 10 presents a comparison of our model against the three baseline models.

As shown in Table 9, our model achieved mean scores of 10.0 for sentence segmentation, 10.0 for logical judgment, 9.7 for translation, and 9.3 for summary and induction, with low variance across the three test epitaphs. These results indicate that the model performs reliably across different epitaph texts, with particularly strong performance in sentence segmentation and logical judgment. The slightly lower scores in translation and summary tasks reflect the inherent complexity of rendering Classical Chinese into fluent modern Chinese while preserving domain-specific nuances.

Capability type	Epitaph Cui Yi	Epitaph Liu Yi	Epitaph Lihua	Mean score	Variance
Sentence breaking	10	10	10	10	0
Logical judgment	10	10	10	10	0
Translation	10	9	10	9.7	0.2
Summary and induction	10	9	9	9.3	0.2

Table 9. Ability evaluation score of epitaph model (Full Score: 10 points)

Table 10 reveals a substantial performance gap between our domain-specific model and the general-purpose baselines. Our model achieved an overall mean score of 9.75, compared to 8.375 for Doubao (base), 8.475 for Qianwen, and 9.15 for ChatGPT. The largest performance gains were observed in sentence segmentation and domain terminology handling, tasks that require deep familiarity with Classical Chinese syntax and Tang dynasty epitaph conventions. These results provide strong empirical support for the core claim of this study: domain-specific fine-tuning is essential for achieving high performance on specialised tasks involving ancient Chinese texts, as general-purpose large language models, despite their broad capabilities, struggle with the unique linguistic and domain-specific demands of epitaph processing.

Model	Sentence segmentation	Logical judgment	Translation	Summary	Overall mean
Doubao (base)	8.2	8.6	8.4	8.3	8.375
Qianwen	8.5	8.3	8.7	8.4	8.475
ChatGPT3.5	9.1	9.0	9.2	9.3	9.15
Ours (Epitaph Model)	10	10	9.7	9.3	9.75

Table 10. Comparison with baseline models (Full Score: 10 points)

System limitations

The evaluation process identified several limitations that should inform future improvements to the model. First, the model exhibited inconsistency in retrieving burial date information: when querying for two individuals with identical burial dates, the model returned discrepancies in one case. Second, while the model correctly understood the classical term *qianhua* (迁化, a euphemism for death), it failed to retrieve relevant cases containing this term when queried. Third, in certain instances, burial date queries returned no information, yet requesting the full text of the same epitaph yielded accurate burial dates; subsequent queries then performed correctly, suggesting potential issues with retrieval consistency. Finally, when queried for an epitaph not present in the knowledge base, the model returned an unrelated epitaph; upon further correction, it fabricated a non-existent epitaph.

These limitations highlight areas for future research, including improving retrieval consistency, enhancing handling of low-frequency terms, and incorporating robust mechanisms to mitigate hallucination.

Optimisation

To address the issues identified during the evaluation phase, targeted optimisations were implemented. When the model encounters questions that it cannot answer or lacks relevant information, it is configured to inform the user truthfully, as per the prompt settings, rather than fabricating responses.

In addition, future research will involve inputting dictionary-type books related to epitaphs into the model to deepen its understanding of epitaph content and enhance the accuracy of information extraction. These resources include the following:

- *Yuanhe xingzuan* and *Gujin xingshi Shu bianzheng*. These works provide comprehensive information on surnames and clan origins, essential for understanding the lineage and historical background of individuals mentioned in epitaphs.
- *Tang cishi kao quanbian* and *Tang pushao chenglang biao*. These records offer details of official titles and administrative divisions during the Tang dynasty, aiding in the interpretation of official positions and titles found in epitaphs.
- *Tang liangjing chengfang kao* and *Sui tang liangjing fangli pu*. These resources offer insights into the geographical and administrative divisions of the Tang and Sui dynasties, which are valuable for understanding the locations and regional contexts mentioned in epitaphs.

Advantages and existing issues of the model

The developed large language models for epitaphs offer several advantages. First, it provides comprehensive interpretation and information extraction. The model facilitates intelligent indexing, reading comprehension, automatic punctuation, annotation, translation, information extraction, annotation generation, social network construction, and biography generation from epitaph texts. Second, the model can handle fragmentation in epitaphs. As many original epitaph

inscriptions are fragmented owing to aging and missing content in collected rubbings, the model assists in reconstructing and completing fragmented texts.

However, certain issues remain. The first concerns the accuracy of information extraction. The model occasionally generates inconsistent or incorrect information, such as mismatched burial dates or omissions of key terms such as ‘迁化’. The second concerns the handling unavailable information. When the model encounters questions that it cannot answer or lacks relevant information, it sometimes fabricates responses instead of acknowledging a lack of data.

Conclusion and prospects

Large language models have the potential to revolutionise the preservation, interpretation, and use of ancient library texts, presenting formidable challenges and new development opportunities. Libraries urgently need to reconstruct new conversational intelligent knowledge service systems based on artificial intelligence by leveraging their high-quality ancient text resources. By actively developing domain-specific models tailored to various genres and research areas within ancient texts, libraries can provide new knowledge services for the use of ancient resources. Furthermore, integrating these large models into teaching and research scenarios can stimulate students' interest and creativity, thereby promoting the dissemination and inheritance of traditional cultures.

Selection of large language models for ancient texts

Developing large language models, especially those focused on ancient texts, requires extensive high-quality, specialised data resources and incurs significant training costs. For public institutions such as libraries, independently developing and deploying foundational large language models may not be the most feasible approach, owing to the associated costs and technical complexities. The best strategy is to leverage the library's various trusted disciplinary resources by collaborating with technology companies to co-develop specialised large language models in the ancient text domain. Alternatively, fine-tuning and optimising open-source foundational large language models can create application scenarios tailored to library services in ancient text fields, instead of building models from scratch.

Construction of social relationship analysis and visualisation of character data

In the large model corpus, biographical data, from resources including the [China Biographical Database](#), [Database of Candidates for Imperial Examinations in Previous Dynasties](#), official dictionaries, and genealogies have been added. A large number of digital resources of ancient books, such as the existing *Siku Quanshu* and the Chinese Basic Ancient Books Database, have also been added to strengthen the marking of the logical relationships between people's lineage, time, and official positions. Combined with natural language processing technology and historical sociology, a more complete ancient interpersonal space-time network has been presented to help digital humanities research.

Building scientific multimodal digital content and establishing multimodal intelligent retrieval

With the ongoing iteration and optimisation of large models, such as GPT-4 and DeepSeek, it is now possible to process both text and digitised images of ancient rubbings and manuscripts. This advance indicates that both text and images can be efficiently converted into processable data within a large model framework.

In this context, librarians can use their interdisciplinary expertise to optimise the annotation of ancient text data knowledge elements and construct a comprehensive knowledge system encompassing multimodal ancient digital text content. By developing comprehensive retrieval capabilities across modalities, including images, text, and videos, they can overcome the

limitations of text-based retrieval. In addition, this approach enables the automatic construction of intelligent knowledge network graphs with entity nodes and relational edges, facilitating cross-domain integration, multimodal compatibility, and conversational interaction functions. However, this transformation requires a profound shift from traditional text-based retrieval to a comprehensive knowledge service platform.

Collaborative construction and sharing of high-quality ancient text training corpora

Several public ancient text databases, such as the [Shidian Ancient Texts](#), [Database of Chinese Epitaphs through the Ages](#), [Tang and Song Literature Chronological Maps](#), [National Ancient Books Census Registration Basic Database](#), [Chinese Ancient Books Bibliography Database](#), and [Chinese Ancient Books Resource Library](#), are available. However, these databases primarily meet basic retrieval needs and have simple applications. Consequently, when used for training high-quality large language models, these databases still fall short in terms of both quantity and quality, particularly owing to the lack of meticulously processed and contextually annotated data.

The unique characteristics of ancient texts, such as the use of Classical Chinese, the frequent occurrence of variant characters, and rich metaphoric allusions, impose stricter requirements on the semantic understanding capabilities of large language models. Therefore, the demand for professionally annotated and refined corpora has increased. To address this challenge, relevant institutions and professionals in the ancient text field must adhere to principles such as data partitioning and open access to accelerate the establishment of collaborative and shared mechanisms for corpus construction. This initiative aims to provide reliable training corpora for large language models in the ancient text domain.

In summary, integrating large language models into ancient text studies to generate vertical domain models or intelligent agents holds immense potential for revitalising, interpreting, and expanding access to ancient resources in libraries. By cleaning and proofreading the underlying data of epitaphs, images can be transformed into high-quality text. Through deep knowledge extraction and graph construction, text can be transformed into structured data. With the help of knowledge graphs and vector databases, libraries can provide intelligent services for digital humanities researchers and carry out popularisation services for the general public and interdisciplinary readers, thus contributing to the protection and dissemination of traditional culture.

About the authors

Yu Minjie is Deputy Director of Zhejiang University Library and a PhD student at the College of Education, Zhejiang University. His interdisciplinary research draws from information science and educational technology, focusing on issues including the construction of large-scale models for vertical domains. He can be contacted at mjyu@zju.edu.cn

Jin Haoyue is a lecturer at Taizhou University. Her research focuses on intelligent education and the integration of information technology with Chinese subject teaching.

Xue Shiliang is a librarian of Zhejiang University Library. His work involves the collation, cataloguing, and intelligent use of ancient Chinese books.

References

- Brzustowicz, R. (2023). From ChatGPT to CatGPT: The implications of artificial intelligence on library cataloging. *Information Technology and Libraries*, 42(3). <https://doi.org/10.5860/ital.v42i3.16295>
- Central Committee of the Communist Party of China, & General Office of the State Council. (2022). *Opinions on promoting ancient books work in the new era*. https://www.gov.cn/gongbao/content/2022/content_5687500.htm (Archived at https://web.archive.org/web/20250901180020/https://www.gov.cn/gongbao/content/2022/content_5687500.htm)
- Chen H., Chen, H., Han, K., Zhu, G., Zhao, Y., & Du, Y. (2024). Customization of vertical domain large models: Theoretical basis and key technologies. *Journal of Data Acquisition and Processing*, 39(3), 524–546. <https://doi.org/10.16337/j.1004-9037.2024.03.003>
- Chen, L. (2022). Issues in the digitization of ancient books and the construction of classical knowledge bases from the perspective of digital humanities. *Journal of Library Science in China*, 48(2), 36–46. <https://doi.org/10.13530/j.cnki.jlis.2022013>
- Guo, Y., Kou, X., Feng, S., & Li, S. (2024). Empowering library reference services with large language models: Logic, scenarios, and systems. *Library Tribune*, 1–10.
- Hu, H. F., & Tang, Y. (2013). *Catalogue of epitaph rubbings from various dynasties in Peking University Library*. Shanghai Chinese Classics Publishing House.
- Huang, Q., Tao, M., Zhang, C., An, Z., Jiang, C., Chen, Z., Wu, Z., & Feng, Y. (2023). Lawyer LLaMA technical report. *arXiv*. <https://doi.org/10.48550/arXiv.2305.15062>
- Lappalainen, Y., & Narayanan, N. (2023). Aisha: A custom AI library chatbot using the ChatGPT API. *Journal of Web Librarianship*, 17, 37–58. <https://doi.org/10.1080/19322909.2023.2221477>
- Li, H. Y. (2019). *Research on restoration methods of digital epitaph rubbings* [Master's thesis, China West Normal University].
- Li, S., Hu, R., & Wang, L. (2024). Construction and application research of ancient Chinese large language models. *Language Strategy Research*, 9(5), 22–33.
- Luo, Y., Yang, K., Hong, M., Liu, X. Y., & Nie, Z. (2023). MolFM: A multimodal molecular foundation model. *arXiv*. <https://doi.org/10.48550/arXiv.2307.09484>
- Luo, Y., Zhang, J., Fan, S., Yang, K., Hong, M., Wu, Y., Qiao, M., & Nie, Z. (2024). BiomedGPT: An open multimodal large language model for biomedicine. *IEEE Journal of Biomedical and Health Informatics*. <https://doi.org/10.1109/JBHI.2024.3505955>
- Rong, L. (2024). Research on prompt learning for automatic classification of books using large language models. *Library Science Research*, (1), 86–103. <https://doi.org/10.15941/j.cnki.issn1001-0424.2024.01.012>
- Song, Z., Yan, B., Liu, Y., Fang, M., Li, M., Yan, R., & Chen, X. (2025). Injecting domain-specific knowledge into large language models: A comprehensive survey. In *Findings of the Association for Computational Linguistics: EMNLP 2025* (pp. 25297–25311). Association for Computational Linguistics. <https://doi.org/10.18653/v1/2025.findings-emnlp.1379>
- Verma, M. (2023). Novel study on AI-based chatbot (ChatGPT) impacts on the traditional library management. *International Journal of Trend in Scientific Research and Development*, 7(1),

- 961–964. <https://bit.ly/4xdGtP4> (Archived by the Internet Archive at <https://bit.ly/3UZWgni>)
- Wang, D. (2023). XunziALLM [Large language model]. GitHub. <https://github.com/Xunzi-LLM-of-Chinese-classics/XunziALLM>
- Wu, N., Liu, C., Liu, J. F., & Wang, D. B. (2024). AIGC-driven automatic summarization of ancient texts: From natural language understanding to generation. *Library Tribune*, 44(9), 111–123.
- Wu, X., Yan, D., & Cai, Y. (2024). Innovative practice of AI literacy education based on the ARCS model. *Library Journal*, 1–12. <https://doi.org/10.13663/j.cnki.lj.2024.11.006>
- Wang, X. (n.d.). Mu ming ju li [Examples of epitaph compositions]. In *Complete library of the four treasures* (Wen-Yuan-Ge ed., Vol. 1482). Shanghai Classics Publishing House. (Original work published during the Ming dynasty).
- Xiao, T. (2025). *Wenxuan* [Anthology of literature] (Facsimile reprint). Chongwen Publishing House. (Original work published ca. 520–530 C.E.)
- Xu, Q. K., Wang, D. B., Liu, Y. T., & Huang, S. Q. (2024). Research on the construction of ancient books knowledge base for knowledge services. *Information Science*, 42(5), 149–158. <https://doi.org/10.13833/j.issn.1007-7634.2024.05.016>
- Zhang, H., Chen, J., Jiang, F., Yu, F., Chen, Z., Li, J., Chen, G., Wu, X., Zhang, Z., Xiao, Q., Wan, X., Wang, B., & Li, H. (2023). HuatuoGPT, towards taming language model to be a doctor. *arXiv*. <https://doi.org/10.18653/v1/2023.findings-emnlp.725>
- Zhang, J. D., Yang, S. H., Liu, J. F., & Huang, Q. (2024). Empowering traditional Chinese medicine ancient texts with AIGC: Construction of the Huang-Di large model. *Library Tribune*, 44(10), 103–112.
- Zhang, X., Tian, C., Yang, X., Chen, L., Li, Z., & Petzold, L. R. (2023). AlpaCare: Instruction-tuned large language models for medical application. *arXiv*. <https://doi.org/10.48550/arXiv.2310.14558>
- Zhao, S., Zhang, J., & Nie, Z. (2023). Large-scale cell representation learning via divide-and-conquer contrastive learning. *arXiv*. <https://doi.org/10.48550/arXiv.2306.04371>
- Zhao, S. Q., Zhou, Y., Ren, Y., Chen, Z., Jia, C., Fang, Z., Long, Z., Liu, S., & Lan, M. (2025). Fùxì: A benchmark for evaluating language models on ancient Chinese text understanding and generation. *arXiv*. <https://doi.org/10.48550/arxiv.2503.15837>
- Zhao, Y. (2018). Preliminary discussion on the digital processing of Tang Dynasty epitaphs. *Library Tribune*, 38(6), 49–55.
- Zhao, Z. X., Hu, D., Liu, C., Shen, S., & Wang, D. B. (2024). Performance evaluation of Chinese general large models in humanities and social sciences. *Library and Information Service*, 68(13), 132–143. <https://doi.org/10.13266/j.issn.0252-3116.2024.13.012>
- Zuo, L., Zhao, Z., & Wang, D. (2024). Automatic classification of "Siku Quanshu" based on large language models. *Journal of Information Resources Management*, 14(5), 22–35. <https://doi.org/10.13365/j.jirm.2024.05.023>

Copyright

Authors contributing to *Information Research* agree to publish their articles under a [Creative Commons CC BY-NC 4.0 license](https://creativecommons.org/licenses/by-nc/4.0/), which gives third parties the right to copy and redistribute the material in any medium or format. It also gives third parties the right to remix, transform and build upon the material for any purpose, except commercial, on the condition that clear acknowledgment is given to the author(s) of the work, that a link to the license is provided and that it is made clear if changes have been made to the work. This must be done in a reasonable manner, and must not imply that the licensor endorses the use of the work by third parties. The author(s) retain copyright to the work. You can also read more at: <https://publicera.kb.se/ir/openaccess>