



An Analysis of the Application of Large Language Models in the Construction of Smart Libraries

Wenping Sun

Zhejiang Financial College, Hangzhou, China
Email: 15869132869@163.com

How to cite this paper: Sun, W.P. (2025) An Analysis of the Application of Large Language Models in the Construction of Smart Libraries. *Open Access Library Journal*, 12: e13204. <https://doi.org/10.4236/oalib.1113204>

Received: March 4, 2025

Accepted: April 21, 2025

Published: April 24, 2025

Copyright © 2025 by author(s) and Open Access Library Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

Large language models (LLMs) exhibit a high degree of compatibility with the development of intelligent libraries, finding applications in various scenarios, including intelligent services for readers, intelligent assistance for librarians, and intelligent construction of digital twins. This paper uses the Bookan LLM as an example to analyze its construction approach and practical application effects. Subsequently, it proposes several strategies for advancement, such as strengthening top-level design, increasing technology research and development, improving management systems, deepening industry-academia-research cooperation, and cultivating professional talents. These strategies aim to promote the healthy and sustainable development of LLMs in constructing intelligent libraries, thereby further enhancing the service level and management efficiency of libraries.

Subject Areas

Library, Intelligence and Philology

Keywords

Large Language Models, AIGC Technology, Artificial Intelligence, Smart Library

1. Introduction

The rapid advancement of AIGC technology, particularly with Large Language Models (LLMs) like ChatGPT, has demonstrated significant application potential across various domains. The integration of LLMs holds revolutionary implications for the development of smart libraries. Both public and university libraries

are tasked with meeting the public's increasing demand for cultural information. They should proactively seize the historic opportunities presented by the information technology revolution, strategically utilize LLMs, and seek innovations in smart library services to provide more convenient and efficient cultural services. This paper examines the compatibility of LLMs with smart library development from a foundational logic and technical perspective. It explores the application scenarios of LLMs in smart library construction, including reader services, librarian assistance, and digital twins. Subsequently, it uses the Bookan LLM as a case study to discuss the practical application of LLMs and derive promotion strategies for smart library development, aiming to provide a reference for the intelligent transformation of libraries.

2. Adaptability of Large Language Models for the Construction of Intelligent Libraries

The evolution of Large Language Models (LLMs) can be traced back to early language models of the 1980s, such as n-gram models, which primarily relied on statistical methods to predict the next word in a sequence [1]. Over time, these models gradually became more sophisticated and advanced. The early 21st century witnessed a significant increase in computational power and the emergence of big data technologies, which paved the way for the rise of neural network-based language models. In particular, the introduction of the Transformer architecture in 2017 marked a revolutionary shift in LLMs, with its superior performance in handling long-range dependencies significantly enhancing the models' ability to understand and generate language. Subsequently, the advent of the GPT (Generative Pre-trained Transformer) series of models signaled a major advancement in natural language processing. These models, through pre-training on massive text datasets followed by fine-tuning for specific tasks, have achieved high performance across various language tasks [2].

From a fundamental perspective, LLMs are natural language processing (NLP) models rooted in deep learning. They acquire the capacity to comprehend and generate human language through training on extensive textual datasets. LLMs typically employ large-scale neural network architectures, utilizing multi-layered non-linear transformations to capture the complexity and diversity inherent in language. These models have demonstrated exceptional performance in various NLP tasks, including text generation, translation, summarization, and question-answering. Libraries, as crucial repositories and disseminators of culture, primarily serve diverse reader groups. The construction of smart libraries aims to provide more personalized, precise, and user-friendly services, guided by user needs. This objective aligns with the development philosophy and application scope of LLMs [3]. The application of LLMs in smart library services can be broadly categorized into the following steps: 1) gathering foundational information, 2) natural language processing, and 3) enhancing generalization capabilities. Technically, LLMs possess robust NLP capabilities, offering the following technological support for smart libraries: 1) Text Mining: Automating the processing of the library's vast

literature collection to extract key information and provide personalized recommendations to users. 2) Intelligent Question Answering: Developing a question-answering system to offer real-time, accurate consultation services [4]. 3) Sentiment Analysis: Analyzing user reviews and feedback to understand user needs and optimize library services. Consequently, LLMs exhibit a high degree of compatibility with the development of smart libraries.

3. Key Applications of Large Language Models in the Development of Smart Libraries

3.1. Reader-Oriented Smart Services

Large language models (LLMs) are pivotal in reader services within smart library environments. They function as virtual assistants, offering 24/7 intelligent question-answering and consultation services. Leveraging semantic understanding, LLMs facilitate efficient information retrieval and personalized recommendations. Integrated with voice assistant systems, they enable voice-based navigation, search, and reading services. Furthermore, LLMs extend to additional services, such as assisting readers with literature review and academic writing. By analyzing reader data, including borrowing records and search behaviors, LLMs provide insights into reading interests, needs, and behavioral patterns. These analyses inform resource allocation and service strategy optimization within the library. Moreover, LLMs possess multilingual processing capabilities, providing translation and cross-lingual search services for non-native speakers, thereby reducing language barriers. This comprehensive approach enhances the quality of library services, meets reader demands, and fosters innovation in library service development.

3.2. Intelligent Assistance for Library Professionals

The intelligent assistance provided by large language models (LLMs) for library staff encompasses the automation of routine administrative tasks, intelligent classification and indexing of library materials, and the provision of real-time data analysis support. The automation of daily administrative tasks includes the automated response to frequently asked questions, and processing book returns and renewals via LLMs, which can significantly reduce repetitive tasks for library staff, thereby alleviating their workload. The intelligent classification and indexing of library materials leverages the data processing capabilities of LLMs to rapidly and accurately classify and index books, thereby improving the efficiency of book retrieval. Real-time data analysis support can assist library staff in better understanding the real-time operational status of the library; for example, by analyzing borrowing data to optimize book procurement and collection layout. These application scenarios not only improve the management efficiency of the library but also provide library staff with powerful decision-making support tools.

3.3. Intelligent Construction for Digital Twins

Digital twins represent a sophisticated technology that employs digital methods

to create comprehensive and precise models of physical entities or systems. Large language models (LLMs), through data integration, can facilitate more accurate and dynamic library management, thereby supporting the development of digital twins for smart libraries. The digital twin model provides real-time feedback on the utilization of the library's physical space, while LLMs can indirectly gather data on user traffic, predict peak periods, and optimize space layout and resource allocation. Furthermore, the integration of digital twins with LLMs allows for the simulation of various service scenarios, enabling the evaluation and adjustment of library operational strategies to meet evolving reader needs. For instance, LLMs can generate virtual guides to offer users an immersive experience, and when combined with the Internet of Things (IoT), they can enhance the library's security monitoring and early warning systems. Through these applications, libraries can provide more intelligent and personalized services, thereby improving overall operational efficiency and reader satisfaction.

4. An Initial Exploration of Large Language Model Applications: A Case Study of the Bookan Large Language Model

4.1. Book a Large Language Model

On September 15, 2023, the inaugural release of the Bookan ChatBK1.0 intelligent consultation cloud application occurred during the "New Ecology and Transcendence: Co-creating the Intelligent Future of Libraries" thematic sub-forum at the China Library Annual Conference, injecting fresh impetus into the intelligent construction of libraries [5]. Unlike other prevalent general-purpose LLMs, the Bookan large language model is specifically designed to address the challenges and pain points currently faced by the development of the library sector, offering enhanced specialization and superior service scalability within the library domain. The question-answering algorithm of the Bookan large language model has been registered and processed within the Internet Information Service Algorithm Filing System and has secured commercial licenses from Baichuan Intelligent, Zhipu, and CPM-Bee large models. This establishes a solid foundation for delivering more efficient and secure internet information services, signifying the recognition of the Bookan large language model's technological prowess in the field of artificial intelligence.

4.2. Core Construction Principles

The primary development strategy for the Bookan large language model involves leveraging artificial intelligence and big data algorithms to furnish libraries with a comprehensive suite of online and offline intelligent interaction platforms. This approach aims to align the intelligent services offered by libraries more closely with user expectations, thereby enhancing user satisfaction with smart libraries. Ultimately, this initiative seeks to comprehensively advance resource development, service delivery, and management within the library system.

4.2.1. Unified Platform

The Bookan large language model integrates the library's internal operational systems and the vast digital resources within the Bookan database into a unified platform. This integration provides patrons with precise search results and relevant library information, significantly enhancing the efficiency and accuracy of information retrieval. Through the model's backend, library staff can regularly enrich and update the library's knowledge base, while patrons can query and provide feedback in real-time via the LLM's frontend. The platform automatically handles general user inquiries around the clock, freeing up staff to focus on more complex issues. Furthermore, the unified platform reduces model training time and simplifies system operation, resulting in more intelligent and efficient services.

4.2.2. Smart Enhancement

Models with a larger number of parameters typically exhibit enhanced expressive power and generalization capabilities. However, this does not necessarily translate to superior performance in real-world applications within a library setting compared to models with fewer parameters. The selection of an appropriate model for practical implementation necessitates the consideration of various factors, including the nature of the task, the volume of available data, and computational resource constraints. In contrast to general-purpose LLMs designed to address a broad spectrum of problems, vertical LLMs are more focused on specific domains. These models are trained with a focus on more targeted data and features, often demonstrating greater sensitivity to data, and can learn effectively with a smaller dataset. When performance is comparable in a specific domain, vertical LLMs offer higher computational efficiency and lower hardware resource consumption, thus achieving a better balance between effectiveness and cost-effectiveness. The Bookan large language model is a vertical model specifically designed for the library domain. Its data sources encompass both broad and specialized information, enabling its natural language understanding capabilities and interaction formats to be comparable to those of general-purpose LLMs such as ChatGPT. Furthermore, it demonstrates superior performance in addressing professional inquiries specific to the library field or questions related to a particular library, providing more accurate responses to user queries within a library context.

4.2.3. Online and Offline Integration

The Bookan large language model (LLM) leverages artificial intelligence (AI) plus software plus hardware paradigm to furnish a unified online and offline consulting service platform for the intelligent construction of libraries, encompassing the entire library and extending to online reader users. The Bookan LLM offers an application programming interface (API) that can be readily integrated into existing systems. Furthermore, it is equipped with digital human displays, mobile terminals, and computer-based user interfaces, enabling users to engage in information consultation and intelligent interaction anytime, anywhere. The service scenarios are illustrated in **Figure 1**.

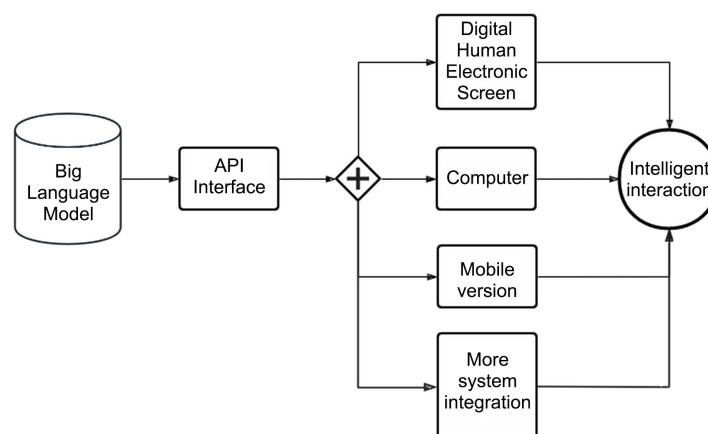


Figure 1. Bookan large language model service scenarios.

4.2.4. Data Processing

The operational steps of the Bookan large language model (LLM) during the data processing phase are similar to those of general LLMs, primarily encompassing the following stages: 1) Data Collection: Data is gathered from diverse sources, including websites, databases, and external APIs. Furthermore, due to the specific requirements of each library, which operates the LLM independently, each library is required to upload its relevant data to a locally deployed backend; this step is managed by specialized librarians. 2) Data Cleaning: The collected data undergoes cleaning to eliminate unnecessary information, duplicate entries, and erroneous data. 3) Data Preprocessing: Preprocessing of the cleaned data involves tokenization, stop word removal, part-of-speech tagging, and named entity recognition. 4) Data Vectorization: Data is vectorized and loaded into a vector database upon system initialization. 5) Intent Completion: User intent is completed after intent recognition and multi-turn dialogue, enabling the execution of various tasks such as QA question answering, internal system retrieval, and external API calls. 6) Reasoning and Merging: Data with high similarity is submitted to the LLM for reasoning and merging. 7) Safety Filtering: The LLM-generated responses undergo safety checks and filtering before final output. The specific steps are illustrated in **Figure 2**.

4.2.5. Security Strategies

During the corpus training phase, the Bookan large language model selects training data aligned with core socialist values to ensure ideological control over the output information. In the user question-answering phase, the model adheres to the principle of generating prompts within the scope of the existing corpus text content to ensure the safety and controllability of the output information. In the generative response filtering phase, a keyword blacklist and other manual intervention mechanisms are established in the background, and open-source online sensitive word libraries are utilized to filter sensitive words generated by the large language model to address real-time, emergent needs. Furthermore, the system supports local deployment, preventing the uploading of internal library materials

to the internet and thus safeguarding user information privacy. The security strategies for each phase are illustrated in **Figure 3**.

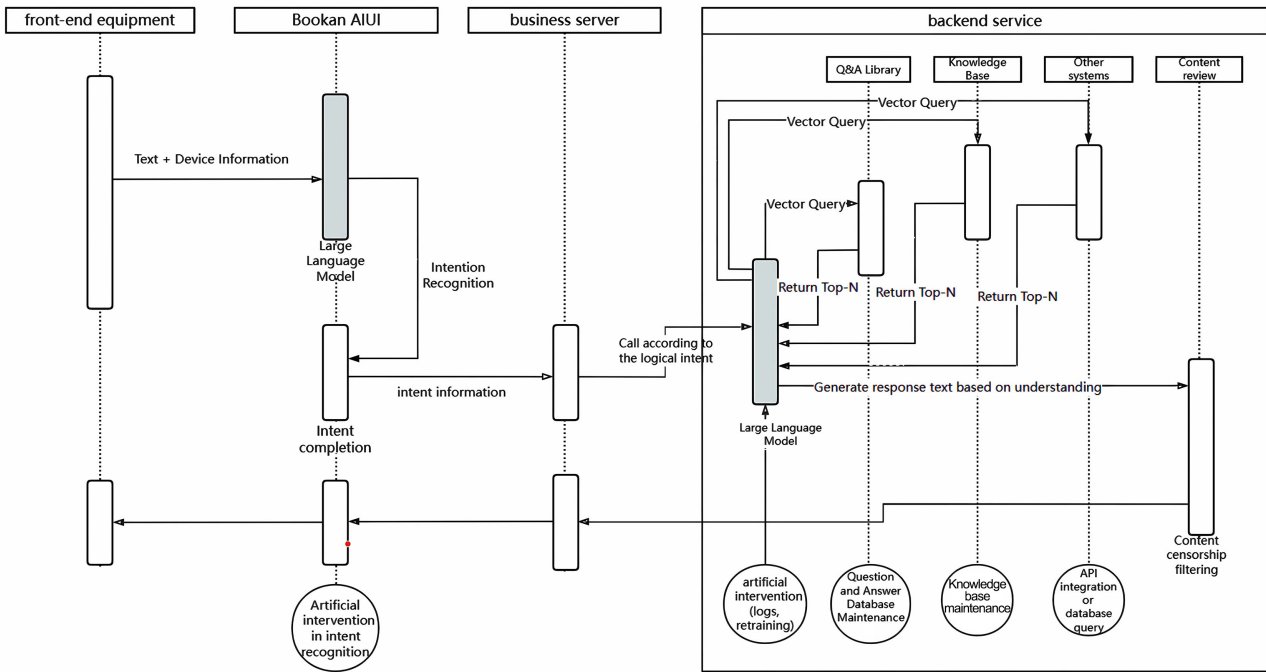


Figure 2. Data processing steps for the bookan large language model.

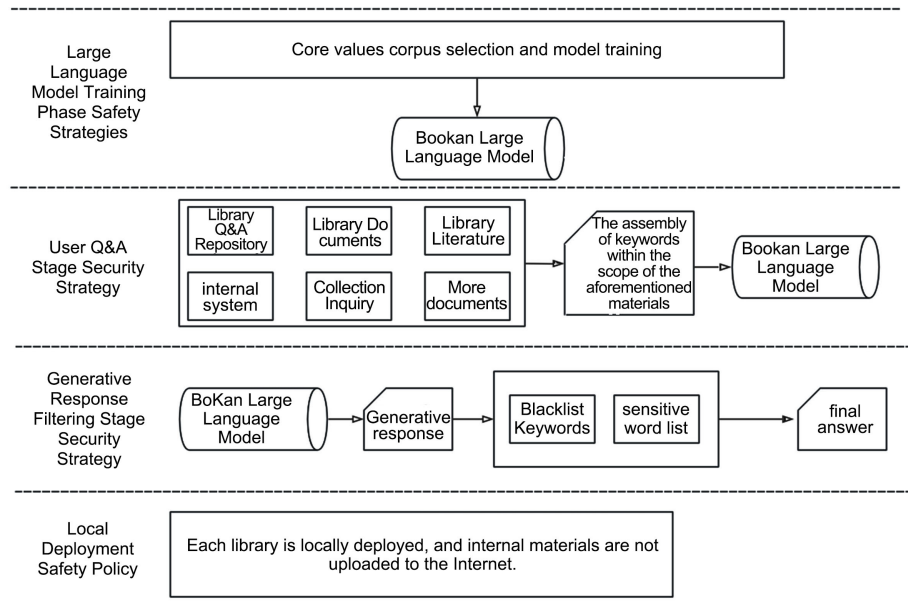


Figure 3. Security strategies for the bookan large language model across each phase.

4.3. Application Effect Evaluation: A Case Study of Zhejiang Financial College

The Bookan large language model (LLM) has been implemented in select libraries. The Zhejiang Financial College Library initiated the integration of this model by

first tasking specialized librarians with the consolidation of the library's extensive proprietary textual data. This data was subsequently uploaded to the LLM's backend for training, thereby imbuing the model with the library's unique characteristics. Consequently, patrons can now access rapid and reliable information consultation services through interactive dialogues. Beyond the conventional question-and-answer interfaces available on web and mobile platforms, the library's main hall features a digital human electronic display, a distinctive feature of the Bookan LLM, which supports 3D high-fidelity, visually realistic virtual human services. The virtual librarian, presented on the electronic screen, exhibits a range of lifelike animations, including blinking, smiling, lip-syncing, waving, and bowing, and can engage in casual conversations with patrons upon their approach. Through natural language input and output, both voice and text-based, the virtual human facilitates a variety of services, including information retrieval, literature analysis and recommendations, and thematic content discussions. It also enables patrons to quickly access information regarding current operational policies, service procedures, and resource utilization methods. This innovation has garnered significant attention from patrons within a short period, quickly becoming the most popular intelligent device in the library.

Since the implementation of the large language model (LLM) within the library, several areas for improvement have emerged: 1) Service content limitations. As a newly introduced intelligent initiative in the library, the LLM's positioning requires clarification, and promotional efforts need enhancement. Currently, a significant number of patrons engage in basic conversations out of curiosity, with the most utilized function still being online e-book reading. The LLM's powerful natural language interaction capabilities require further demonstration. 2) Inadequate management system. Despite the LLM's excellent performance in intelligent interaction, it struggles with daily maintenance and updates. Furthermore, the absence of an effective feedback mechanism prevents timely resolution of issues encountered by users, thereby impacting user experience [6]. 3) Staff proficiency needs improvement. Apart from the specialized librarians responsible for this service, other staff members exhibit varying levels of proficiency with the LLM, which directly affects the model's utilization efficiency and the patrons' experience.

5. Strategies for Advancing the Application of Large Language Models in Smart Library Development

5.1. Strengthening the Top-Level Design

LLMs offer significant potential to enhance library service quality, reader experience, and management efficiency. Therefore, the overarching design should clearly define the development goals of the smart library, integrating the application of LLMs with the library's long-term development strategy. The role of LLMs in the construction of smart libraries is to assist libraries in achieving intelligent services, with specific application scenarios including intelligent question answer-

ing, book recommendations, and knowledge services. Libraries should formulate a comprehensive application plan for LLMs in smart libraries as early as possible, selecting LLMs that are suitable for the library's characteristics, and rationally allocating resources such as human resources, material resources, and financial resources to promote the phased implementation of the project. In the top-level design, libraries should also strive to obtain support and guarantees from national and local governments in terms of policies and funding. Simultaneously, they must comply with relevant laws and regulations to ensure that the application of LLMs is compliant and legal, thereby creating favorable conditions for the orderly and efficient application of LLMs in the construction of smart libraries.

5.2. Intensifying Technological R&D

LLMs in the construction of intelligent libraries remain a cutting-edge technical challenge, and increased technological research and development (R&D) is crucial for promoting its widespread implementation. Libraries should establish deep collaborative relationships with research institutions, enterprises, and other libraries to form a cross-disciplinary R&D team. This team should include artificial intelligence experts, library science specialists, software engineers, and data scientists. They should identify key areas for technology R&D based on the specific needs of the library, such as intelligent question answering, personalized recommendations, and knowledge graph construction. The R&D team should establish a series of technology R&D projects, conduct specialized research for different application scenarios, and accelerate the technology R&D process by driving theory and practice through these projects. Libraries should also construct a high-quality library domain dataset to support the training and testing of large language models [7]. In the future, LLMs should strive to provide readers with more personalized services, recommending relevant library resources based on the interests of different readers or predicting their potential needs based on their consultation history, thereby providing more accurate services.

5.3. Establishing a Robust Management Framework

A comprehensive management system is essential for LLMs to ensure their efficient, stable, and sustainable operation within the context of smart library applications. The construction of this management system should encompass multiple facets, including but not limited to model maintenance and updates, user privacy protection, data security, and service quality monitoring. Initially, a regular model evaluation mechanism should be established to ensure the model's adaptability to the evolving demands of library services, facilitating timely updates to provide the most current knowledge and information [8]. Furthermore, user privacy protection constitutes a critical component of the management system, necessitating the formulation of stringent data processing and storage protocols to safeguard the security of readers' personal information [9]. In addition, a service quality monitoring mechanism can assist libraries in promptly identifying and resolving issues

within the service, thereby ensuring the continuity of the reader experience. A user feedback mechanism should be established to collect user experiences and suggestions regarding the application of LLMs, enhance user interaction, and improve user satisfaction and loyalty. Through the establishment of a robust management system, the application of LLMs in smart libraries will become more mature and reliable.

5.4. Strengthening Industry-Academia-Research Collaboration

LLMs offer robust support for industry-academia-research (IAR) collaborations in the development of intelligent libraries. LLMs facilitate information integration, intelligent retrieval, and data analysis, thereby providing precise knowledge services to all IAR stakeholders. Establishing university-industry collaboration platforms can foster knowledge exchange and technology transfer between academia and industry. Universities and research institutions can contribute theoretical foundations and innovative thinking, while businesses can offer practical application scenarios and market demands, jointly promoting the development and application of LLM technology. Furthermore, collaborative projects and internship opportunities can cultivate a cohort of interdisciplinary talents proficient in both technology and library operations, laying a solid human resource foundation for the long-term development of intelligent libraries. In the practical implementation of intelligent library construction, LLMs should be leveraged to promote open access to library resources, provide rich information resources for IAR collaborations, and support the construction of collaborative innovation platforms, offering venues and essential tools for communication and cooperation among all IAR parties [10].

5.5. Cultivating Specialized Professionals

To optimize the application of LLMs in the development of intelligent libraries, libraries must accelerate the cultivation of a specialized workforce. In alignment with the developmental needs of intelligent libraries, libraries should promptly formulate professional talent development plans. These plans should encompass specialized skills, data analysis, and artificial intelligence knowledge. Leveraging LLMs to analyze industry trends and forecast future talent requirements will provide data-driven support for these training initiatives [11]. Libraries can enhance their staff's comprehension and application of LLM technology through regular professional training sessions, collaborative learning opportunities, and workshops. Furthermore, academic libraries can collaborate with secondary colleges to offer relevant courses and internship programs, attracting and nurturing more young talent interested in library science and artificial intelligence technologies. In addition, libraries should establish incentive mechanisms to encourage professional librarians to engage in LLM-related research and development, further stimulating their innovative potential and contributing more expertise and resources to the construction of intelligent libraries.

6. Conclusion

Large language models (LLMs), as potent natural language processing tools, not only demonstrate a profound theoretical grasp of linguistic complexity but also exhibit substantial practical potential. Their application in the realm of intelligent library services holds considerable promise, with the capacity to significantly enhance service quality and user experience. However, “a one-size-fits-all model has its limitations in practical application,” implying that despite LLMs being trained to comprehend and generate natural language and to handle a broad spectrum of tasks, they still possess certain constraints in real-world scenarios. The deployment of LLMs is further accompanied by technological, ethical, and sustainability challenges, necessitating collaborative efforts from the library community, technology providers, and society at large to ensure their healthy and sustainable development within the context of intelligent library construction.

Funding

This paper presents one of the research findings from the 2023 general youth project, “Research on the Path of AIGC Technology Empowering Intelligent Library Services in the Metaverse Vision” (2023YB28), funded by the Zhejiang Financial Vocational College’s fundamental research grants.

Conflicts of Interest

The author declares no conflicts of interest.

References

- [1] Liang, J., Zhang, L., Yan, S., *et al.* (2024) Research Progress on Named Entity Recognition Based on Large Language Models. *Computer Science and Exploration*, **18**, 2594-2615.
- [2] Wang, R., Zhang, X., Wang, M., *et al.* (2024) Network Public Opinion Multi-Task Analysis Based on Hybrid Retrieval Enhanced Generation Large Language Model. *Journal of Intelligence*, 1-14. <http://kns.cnki.net/kcms/detail/61.1167.G3.20241212.0931.008.html>
- [3] Guo, Y., Kou, X., Feng, S., *et al.* (2024) Large Language Models Empowering Library Reference and Consultation Services: Logic, Scenarios, and Systems. *Library Tribune*, **45**, 118-127. <http://kns.cnki.net/kcms/detail/44.1306.G2.20240220.0947.004.html>
- [4] Yan, D., Xu, Y., Yu, C., *et al.* (2023) Theoretical Progress, Practical Problems, and Future Prospects of Metaverse Libraries. *Library Journal*, **42**, 4-12, 21.
- [5] Cloud Han Application (2023) “ChatBK1.0 Bokan Wisdom Consultation” Introduction. <https://www.calsp.cn/2023/10/20/bulletin-202310-01/>
- [6] Li, Y., Dong, P. and Li, S. (2023) Research Topics and Future Prospects of Metaverse Libraries Based on Functional Positioning. *Library Theory and Practice*, No. 5, 129-136.
- [7] Liu, Q., Liu, S. and Liu, W. (2023) Application Modes and Data Governance of Large Models in the Field of Library and Information Science. *Library Journal*, **42**, 22-35.
- [8] Wei, M., Luo, K., Tang, M., *et al.* (2024) Application Research of Large Language

- Model (LLM) Embedded Smart Library Scenario Services. *Library Theory and Practice*, No. 1, 83-94. <https://doi.org/10.14064/j.cnki.issn1005-8214.20241029.001>
- [9] Luo, F., Cui, B., Xin, X., *et al.* (2023) Risk Paradigm and Control Strategies of Large Language Models Embedded in Library Knowledge Services. *Books and Information*, No. 3, 99-106.
- [10] Fu, R. and Yang, X. (2023) Analysis of AIGC Language Models and Research on Application Scenarios in University Libraries. *Journal of Agricultural Library and Information Science*, **35**, 27-38.
- [11] Guo, L. and Fu, Y. (2023) Building a Smart Library with Large Language Models: Framework and Future. *Library Journal*, **42**, 22-30, 133.