

ANALYSIS OF THE ARTIFICIAL INTELLIGENCE-BASED “XIAOSHU” RECOMMENDATION SYSTEM IN CHINA

Shakhnoza Rustamova

master of TUIT

+998911866757 shaxnozarustamova229@gmail.com

Abstract: *With the advancement of time and technology, library services and their functions continue to evolve. Globally, library reference services have undergone significant transformations driven by technological progress. Paper-based information sources have gradually transitioned to digital formats, and traditional face-to-face reference consultations have shifted to email-based and online chat services. With the rapid growth of Artificial Intelligence (AI), libraries have entered a new stage of intelligent services. Chatbots that interact through text or voice have enabled the development of smart reference services.*

The application of AI in libraries has not only improved the accuracy of information retrieval and recommendation processes but has also introduced new service models such as personalized reading recommendations, intelligent classification, and knowledge graph construction. In recent years, Large Language Models (LLMs) have matured considerably, and the emergence of ChatGPT-4 has triggered a transformation in library services. These models possess powerful natural language understanding and generation capabilities, significantly enhancing service efficiency and meeting the diverse and personalized needs of modern users [1].

This article examines the challenge of delivering relevant content to users in the context of information overload. The AI-based recommendation system of the Xiaoshu platform is analyzed. The study explores the system’s voice interaction mechanism, knowledge retrieval and integration framework, service experience and user interaction, response accuracy, and overall service effectiveness.

Keywords: *artificial intelligence, recommendation system, Xiaoshu platform, digital library, information retrieval system*

Introduction.

In today’s information age, users face numerous difficulties when searching for the information they need. As the demand for information increases day by day, the volume of available information is also growing rapidly. Finding relevant resources has become a challenge not only in libraries but also in almost every sphere of human life. In order to address these problems and optimize information management, many sectors have undergone

digital transformation. In particular, library information services have digitized their collections and started to utilize recommendation systems.

The first recommendation systems in digital libraries were quite simple, as search engines could only retrieve documents that directly matched specific keywords. The main issue was that traditional search systems could not understand synonyms, distinguish context, or analyze meaning. For example, when a user searches for “artificial intelligence,” the system may fail to retrieve articles about “machine learning” because synonyms are not integrated into the system [2].

A Recommendation System (RS) is an intelligent information system that automatically suggests relevant information, products, or services to users based on their interests, behavior, search history, ratings, and other social and technical indicators [3]. Although such systems were initially developed for commercial purposes (for example, in e-commerce platforms such as Amazon, Netflix, and eBay), their capabilities are now successfully implemented in scientific and educational institutions, including libraries. Recommendation systems play a particularly important role in digital libraries, online learning platforms, and academic research systems. They help better understand user needs, provide relevant resources, and thereby accelerate and improve the efficiency of the information retrieval process [4].

Xiaoshu is a digital twin equipped with voice communication capabilities that imitates the behavior and linguistic style of real librarians, providing natural and human-like interactions. Since its launch at the National Library of China on December 28, 2023, the intelligent librarian “Xiaoshu” has utilized large language models to deliver efficient and accurate information responses, while enhancing user interaction through rich character design [10].

In the future, this research will focus on expanding and updating the database to further improve the functions and services of the intelligent librarian “Xiaoshu” [7]. The main areas of development include database expansion and updates, emotion recognition and response, personalized recommendations, multilingual support, improved response time, and user feedback mechanisms. Through continuous technological innovation and service optimization, the intelligent librarian aims to provide more efficient, intelligent, and user-oriented library services, becoming an integral part of the smart library ecosystem [1].

The purpose of this study is to examine a modern library service model based on generative artificial intelligence and to provide a scientific justification for the possibilities of implementing it in the information and library sector of Uzbekistan.

Today, at the global level, libraries are transforming from their traditional functions of storage and distribution into intelligent information intermediaries. The Xiaoshu system represents a practical example of this transformation, as it not only offers search and recommendation services but also enables emotional and interactive communication with

users. This demonstrates the concept of human-centered AI in the digital library environment [9].

Analysis of the Xiaoshu Platform



Design Diagram.

The physical design of the smart librarian adopts the spatial image of the red public telephone booths that first appeared on the streets of London. This design symbolizes a dedicated space where users can directly interact with the intelligent librarian. The smart librarian is named “Xiaoshu,” meaning “knowledgeable about books and wise in giving advice” [6].

As shown in Figure 1, its appearance represents an AI digital twin of a real librarian. The portrait was also generated by artificial intelligence; however, to avoid the “Uncanny Valley” effect, it is presented in a cartoon style, which also helps attract readers [10].

Inside the telephone booth, there is a desktop equipped with an LCD panel that displays the content of conversations during interactions with the smart librarian. It also presents large volumes of information, web pages, or QR codes when the virtual librarian needs to provide extended responses.

The system consists of five main modules designed to support human–computer interaction between the system and readers. Among them, the “Xiaoshu Playback” module is responsible for instantly animating the facial expressions of the smart librarian while it communicates with users. The voice input and voice output units are responsible for recording and playing audio during interactions.

To give the virtual librarian a more human-like and professional image, it was named “Xiaoshu” (曉書, abbreviated from 通曉圖書), which in Chinese means “knowledgeable about books and wise in offering advice” [10].

Its overall exterior design is inspired by the red British telephone booth, creating a semi-enclosed and immersive communication space. Inside, an LCD display presents textual content and related supplementary information [10].

Xiaoshu is equipped with a camera that detects users standing in front of it and automatically initiates a greeting when someone approaches. Users can press the red button on the right side to start voice interaction [1].

Voice Interaction System.

The user’s voice is transmitted from the front-end system to the back-end STT (Speech-to-Text) service, where speech is converted into text for processing. Simultaneously, the system uses TTS (Text-to-Speech) technology to convert textual responses into synthesized speech [8]. This technology transforms text into natural-sounding audio that imitates the voice of real librarians. The generated voice is then played through the front-end speaker.

The primary goal of these three components is to create an experience in which users interacting with the virtual librarian feel as though they are communicating with a real librarian.

In addition, the “desktop presentation” module is used to display information received from the back-end system on the front-end interface. All content—including conversations between the two parties, recommended book information, QR codes for book collections, popular book lists, new book announcements, library service guidelines, transportation information pages, and nearby location data with Google Maps integration—is displayed on the device’s screen panel located on the desktop, as shown in Figure 1.

The design aims to ensure that Xiaoshu can continue a conversation within 20 seconds, while additional content is presented on the desktop display. This prevents conversations from becoming too lengthy and helps avoid user impatience [7].

Finally, the front-end system also includes a “gender and age group recognition” module. A camera hidden at the front of the telephone booth records the current interaction and the virtual interface. These recordings do not involve personal privacy data. For example, facial recognition is limited to identifying gender and age-related information only, and its purpose is to provide data for future system improvements [10].

Knowledge Retrieval and Generation Mechanism.

In order for Xiaoshu to accurately respond to well-defined questions, NLPI developed four main databases based on common user intentions: library services, books, Xiaoshu’s personal (persona) settings, and information about notable figures.

These databases are integrated with a Retrieval-Augmented Generation (RAG) architecture and the ChatGPT language model, enabling a flexible question-and-answer service based on semantic analysis. Particularly in bibliographic inquiries, this approach highlights the most significant difference between general-purpose generative AI assistants and AI systems specifically adapted to the library environment [10].

The NLPI collection contains more than 1.3 million resources. However, since the size of the database may affect system speed, approximately 550,000 frequently used printed book records were selected as the core dataset. This optimization allows Xiaoshu to respond faster than human reference librarians while ensuring broader knowledge coverage, thereby improving the overall quality of reference services [10].

Service Experience and User Interaction Analysis.

At present, Xiaoshu provides four main service functions:

1. Book recommendations
2. Collection (catalog) search
3. Information about library services
4. Social (open-ended) conversation

Practical observations indicate that, similar to communication with human librarians, social conversation is not only an entry point for engaging users but also an important factor in building trust and long-term interaction. For this reason, NLPI deliberately retained ChatGPT’s conversational capabilities. This encourages users to interact freely and friendly with Xiaoshu, helping them become familiar with the system and motivating regular use of the service [9].

According to system log data from December 1, 2023 (including the testing phase prior to the official launch on December 28) to October 30, 2025, Xiaoshu conducted approximately 70,000 interactions. Assuming 26 working days per month and one request per user, this corresponds to an average of about 117 users served per day [1].

After the first three months of operation, weekend usage exceeded weekday usage, reaching an average of 3,000 interactions per month [10]. Xiaoshu’s personalized design - its self-introduction style, tone modulation, and conversational approach - makes it appear friendly and approachable, similar to a real librarian. As a result, it has become particularly popular among young readers and elderly users.

This trend demonstrates that the AI librarian not only attracts users but also fosters interest and motivation toward libraries and artificial intelligence technologies.

Xiaoshu’s Response Accuracy.

According to the librarians responsible for maintaining and developing Xiaoshu, the system consistently maintains an average monthly accuracy rate of over 80% [10].

Accuracy mainly depends on two factors:

1. Input data quality – including clear voice capture, correct language recognition, and the quality of the speech-to-text (STT) process.
2. Backend database completeness – meaning the availability of relevant responses through the Retrieval-Augmented Generation (RAG) mechanism [5].

Backend accuracy can be improved through regular database updates. However, input data quality is limited by users' speech patterns and the capabilities of generative AI speech models. For example, language support is currently mainly limited to Mandarin and English, while Taiwanese Hokkien is not supported. Additionally, the internal accuracy level of commercial LLM systems (such as ChatGPT-4 and ChatGPT-5) also affects overall performance [6].

Xiaoshu's Service Efficiency.

Currently, NLPI has deployed one Xiaoshu device located in the first-floor lobby of the library, providing real-time voice interaction services. On average, 3,000 interactions occur per month, approximately 30% of which relate to professional (core) library consultations. This corresponds to about 1,000 primary library inquiries handled monthly [10].

If each response takes approximately one minute, this results in about 16.6 hours of librarian labor saved per month [9].

By providing continuous on-site service, Xiaoshu enhances the convenience of library services, reduces staff workload during peak hours, and efficiently handles repetitive inquiries and specific book recommendations. Consequently, human librarians can focus more on tasks requiring technical expertise and human judgment, such as organizing exhibitions, promoting reading programs, conducting information literacy training, and offering professional consultations [1].

Therefore, expanding the number or types of Xiaoshu devices across different floors of the library could further reduce librarians' advisory workload and increase overall service capacity.

Conclusion.

The Xiaoshu experience provides an empirical foundation for examining ethical, technical, and organizational issues related to the application of generative AI technologies in the national information and library system. This can support the development and phased implementation of a "Smart Library" concept within the context of Uzbekistan.

Overall, the analysis of the Xiaoshu system establishes a theoretical and methodological basis for identifying innovative development directions in the national library system, improving service quality, and implementing modern AI-based information services.

References.

1. Xie Xanyi. *Smart Library: An Intelligent Librarian Powered by GPT-4*. Master's Program in Digital Innovation, Tunghai University. – China, 2024. p. 3.

2. Salton G., McGill M. J. Introduction to Modern Information Retrieval. – New York : McGraw-Hill, 1983. – 448 p
3. Ricci, F., Rokach, L., & Shapira, B. (2015). Recommender Systems Handbook (2nd ed.). Springer. <https://doi.org/10.1007/978-1-4899-7637-6>
4. Aggarwal, C. C. (2016). Recommender Systems: The Textbook. Springer. <https://doi.org/10.1007/978-3-319-29659-3>
5. Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., & Kiela, D. (2020). Retrieval-augmented generation for knowledge-intensive NLP tasks. Advances in Neural Information Processing Systems, 33, 9459-9474.
6. Liu, Z.-E. (2024). 聊天機器人之讀者體驗研究：以國立公共資訊圖書館為例 [A study of reader experience with chatbots: A case study of the National Library of Public Information]. In 2024 圖書資訊學術與實務研討會會議論文集 (pp. 199-206). Library Association of the Republic of China (Taiwan).
7. Yang, J., Jin, H., Tang, R., Han, X., Feng, Q., Jiang, H., & Hu, X. (2024). Harnessing the power of LLMs in practice: A survey on ChatGPT and beyond. ACM Transactions on Knowledge Discovery from Data, 18(6), 1–32. <https://doi.org/10.1145/3649506>
8. Zhou, K., Li, J., Tang, T., Wang, X., Hou, Y., & Wen, J. R. (2023). A survey of large language models. arxiv preprint arxiv:2303.18223. <https://arxiv.org/abs/2303.18223>
9. Zheng, Q., Tang, Y., Liu, Y., Liu, W., & Huang, Y. (2022, April). UX research on conversational human-AI interaction: A literature review of the ACM Digital Library. In S. Barbosa & C. Lampe (Eds.), Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems (pp. 1-24). Association for Computing Machinery. <https://doi.org/10.1145/3491102.3501855>
10. Hsiang-Ping Ma, Yun-Fan Chen. Implementation and Application of a Generative AI Virtual Librarian: A Case Study of the National Library of Public Information, Taiwan. - INTERNATIONAL JOURNAL OF LIBRARIANSHIP, 10(4), 137-148