A Study on the Evaluation of AI Literacy Education in University Libraries Based on Multi-Source Data Fusion*

Chu Jiewang^1^ Zhou Keyan^1^ Niu Yue^2^

2026-07-24

ChinaRxiv

Abstract

With the rapid development of GenAI, AI literacy has become one of the core competencies in talent cultivation in higher education. As key venues for learning, university libraries are actively carrying out AI literacy education. However, current scholarship has paid insufficient attention to evaluation mechanisms for AI literacy education, making it difficult to effectively assess AI literacy education conducted by university libraries. From the perspective of university libraries, this paper systematically reviews the evolution and connotations of AI literacy education, analyzes existing problems in current AI literacy education, clarifies the unique advantages of university libraries in carrying out AI literacy education, and introduces the CIPP evaluation model to construct a multi-source data-driven evaluation system. The study effectively maps multi-source data from university libraries onto the four-dimensional indicators of CIPP, and proposes a comprehensive and quantifiable indicator system for evaluating AI literacy education, so as to achieve objective and dynamic assessment of AI literacy education activities in university libraries. Meanwhile, the AHP method is used to determine the weights of indicators at each level, providing theoretical support and practical pathways for university libraries to continuously optimize AI literacy education services.

Full Text

【Special Topic: AI Literacy Education】

A Study on the Evaluation of AI Literacy Education in University Libraries Based on Multi-Source Data Fusion*

Chu Jiewang[1] Zhou Keyan[1] Niu Yue[2]

1. School of Management, Anhui University, Hefei 230601

2. Library of Northwestern Polytechnical University, Xi' an 710072

Abstract With the rapid development of GenAI, AI literacy has become one of the core competencies in the cultivation of university talent. As key sites of learning, university libraries are actively carrying out AI literacy education. However, the current academic community has paid insufficient attention to mechanisms for evaluating AI literacy education, making it difficult to effectively assess the AI literacy education conducted by university libraries. From the perspective of university libraries, this paper systematically reviews the evolution and connotations of AI literacy education, analyzes the problems currently existing in AI literacy education, clarifies the unique advantages of university libraries in carrying out AI literacy education, and introduces the CIPP evaluation model to construct a multi-source-data-driven evaluation system. The study effectively maps multi-source data from university libraries to the four-dimensional indicators of CIPP, and proposes a comprehensive and quantifiable indicator system for evaluating AI literacy education, in order to realize an objective and dynamic evaluation of AI literacy education activities in university libraries. At the same time, the AHP method is used to determine the weights of indicators at all levels, providing theoretical support and practical pathways for university libraries to continuously optimize AI literacy education services.

[**Keywords**] AI literacy; evaluation of AI literacy education; university library; evaluation system; CIPP evaluation model

[**Chinese Library Classification Number**] G252 [**Document Identification Code**] A [**Article Number**] 1003-7845(2026)01-0018-10

[**Citation Format**] Chu Jiewang, Zhou Keyan, Niu Yue. A Study on the Evaluation of AI Literacy Education in University Libraries Based on Multi-Source Data Fusion[J]. *University Library Work*, 2026(1):18-27.

Introduction

Since ChatGPT emerged so dramatically, generative artificial intelligence (Generative Artificial Intelligence, GenAI) has developed rapidly, pressing the “AI+” accelerator for the development of human society. However, while GenAI technology has brought tremendous opportunities, it has also quietly intensified competitive pressure between humans and machines and introduced potential risks such as large-model “hallucinations,” that is, the possible generation of content inconsistent with reality or logically contradictory[1]. Users who lack the corresponding ability to judge knowledge may blindly trust content generated by large models. To ensure that individuals in the era of artificial intelligence possess core competitiveness that will not be completely replaced by AI, and to avoid falling into the traps of knowledge bias or even knowledge misconceptions caused by large models, education aimed at enhancing AI literacy is an important and indispensable component.

As frontier sites for connecting with emerging technologies, universities are able to access AI technologies rapidly and comprehensively. Today, AI has been integrated into every aspect of people’ s lives, and AI literacy has become one of the core competencies that must urgently be mastered in higher education. As information resource centers and academic support platforms, university libraries naturally possess unique advantages in carrying out AI literacy education. On the one hand, university libraries have abundant digital resources and intelligent-technology application environments, providing fertile ground for AI literacy education; on the other hand, university libraries have long undertaken the function of information literacy education and possess the professional capacity to provide related educational services[2]. As an extension of information literacy, AI literacy should likewise be cultivated under the leadership of university libraries.

The key to literacy education lies in establishing a scientific and reasonable evaluation indicator system and concrete practical pathways. An objective and comprehensive evaluation system can test the effectiveness of educational activities, identify omissions and remedy deficiencies in a targeted manner, and provide a decision-making basis for improving literacy education, thereby forming a closed loop for optimizing educational services. However, the current evaluation mechanisms for AI literacy education in university libraries are not yet mature and suffer from many problems. First, current evaluations are mostly focused on educational outcomes—that is, AI literacy levels—while lacking systematic evaluation of key links such as educational context, input, and process. Second, data sources are unitary and lack the capacity to support cross-system, multi-source objective data; there is excessive reliance on traditional questionnaires, tests, and single-dimensional assessments. In response to the above problems, this paper adopts the perspective of university libraries and, by introducing the CIPP evaluation model and combining it with multi-source data generated in the process of AI literacy education activities in university libraries, constructs a multi-source data fusion mechanism for AI literacy education in university libraries, and proposes an evaluation indicator system based on the mapping between the four CIPP dimensions and multi-source data, providing support for university libraries to continuously opti—

* This paper is one of the research outputs of the project "Research on the Cultivation Path of AI + Information Literacy for Young University Teachers" of the Innovation and Development Center for Ideological and Political Work in Higher Education under the Ministry of Education (teacher ideological and political education direction) (Project No.: 2025SZZX0208), and the Anhui University Major Teaching Research Project "Exploration of Intelligent-Driven Personalized Learning Paths" (Project No.: jyxm202504).

Author Profiles: Chu Jiewang, Ph.D., specially appointed professor as a provincial-level leading talent, doctoral supervisor; research interests include knowledge management, innovation management, and science and technology management. Zhou Keyan, master' s student; research interests include knowledge management and science and technology management; corresponding author, Email: 18071358035@163.com. Niu Yue, master' s degree holder, associate research librarian; research interests include smart library construction and artificial intelligence literacy education.

Date Received: 2025-09-07 **Responsible Editor:** Liu Yi

···provide scientific and reasonable practical pathways for AI literacy education services, with the aim of comprehensively improving the quality of AI literacy education and cultivating university talent for society who can adapt to development in the AI era.

1 The Connotation of AI Literacy Education and the Theoretical Basis of CIPP Evaluation

1.1 The Paradigm Shift from Information Literacy to AI Literacy

In the information age, "literacy" has become a rich and dynamically evolving comprehensive concept, and it has taken on different connotations as information technology and modes of production have changed. AI literacy is precisely the inheritance and expansion of information literacy and digital literacy against this background. "Information literacy" was first proposed by P. G. Zurkowski in 1974, who defined it as the ability to solve practical problems through training, mastery of information tools, and acquisition of relevant information[3]. In 2015, the Association of College and Research Libraries (ACRL) interpreted it as a comprehensive ability involving reflective discovery of information, understanding how information is produced and valued, and using information to create new knowledge and participate rationally in learning communities[4]. Synthesizing the views of many scholars, the core content of information literacy includes mastery of basic knowledge such as information concepts, characteristics, and structures; the ability to search for, process, evaluate, and use information; correct recognition of the value of information; a positive attitude; critical thinking ability; and the degree of compliance with information ethics.

In the digital age, information literacy has been upgraded to digital literacy, with greater emphasis on the application of digital tools and methods. The United Nations Educational, Scientific and Cultural Organization (UNESCO) defines it as follows: the ability to safely and appropriately acquire, manage, understand, integrate, communicate, evaluate, and create information through digital technologies. It encompasses computer literacy, information and communication technology literacy, information literacy, media literacy, and other abilities[5], and emphasizes the use of digital technologies to communicate, collaborate, solve problems, and create digital content. Judging from its development trajectory, digital literacy has provided an extremely sufficient foundation for understanding AI in the era of artificial intelligence.

In the era of artificial intelligence, AI literacy is the inheritance and extension of information literacy and digital literacy. It goes beyond nontechnical personnel's simple understanding of AI tools and aims to enable individuals to communicate and collaborate effectively with AI, apply AI tools in real-world scenarios, and maintain a critical evaluative attitude toward AI technologies. The emergence of AI literacy is not merely a simple extension of knowledge and skills; it also represents a comprehensive transformation in how human beings produce, acquire, and use knowledge. Cai Yingchun et al.[6] argue that AI literacy includes both understanding and application of technology and a comprehensive system of skills and knowledge. Xin Bingxin et al.[7] point out that the main elements of AI literacy include knowledge, use, citation, thinking, and ethics.

The *Artificial Intelligence Act* issued by the European Union clearly defines AI literacy as: the skills, knowledge, and understanding that enable providers, deployers, and affected persons, in accordance with their respective rights and obligations under this Act, to make informed deployments of AI systems and to understand the opportunities and risks brought by AI as well as the harms it may cause[8]. In summary, the concept of AI literacy has not yet formed a unified expression in the academic community, and different scholars offer different definitions; however, its connotations are largely consistent. In view of the needs of establishing indicators for the subsequent AI literacy education evaluation system, this paper integrates the views of multiple experts and defines the core elements of AI literacy in the following four aspects: Technical knowledge. This refers to having a certain understanding of basic concepts such as machine learning, natural language processing, and neural networks, and being able to understand and explain the operating principles of AI. Practical skills. This refers to being able to effectively use AI tools, select suitable AI tools in specific contexts, and determine when AI should be used and when human decision-making is needed. Critical thinking. This refers to critically evaluating AI thinking processes, data sources, and data content; maintaining doubts about AI outputs; and reviewing AI-generated content. Ethical awareness. This refers to recognizing that AI must assume ethical responsibilities in the operation of human society, rigorously examining the values and assumptions embedded in AI, and considering potential bias issues.

1.2 The Current Status and Deficiencies of AI Literacy Education

At present, the use of AI tools among university faculty and students is increasingly widespread. According to statistics from third-party organizations, nearly 60% of university faculty and students use GenAI every day or several times per week[9]. Such high-frequency use has improved the efficiency of scientific research and learning, but it has also given rise to misconduct; academic improprieties such as directly adopting AI-generated content and using AI to fabricate research data are becoming increasingly prominent. Although university faculty and students use AI tools at such a high frequency and may even demonstrate a medium-to-high level of AI literacy under certain AI literacy evaluation systems, their critical thinking and ethical awareness have not improved in tandem. It can therefore be seen that current AI literacy education in universities has many shortcomings, and there is an urgent need for a complete educational evaluation system.

In response to the needs of the times, AI literacy education has become a key topic in the current education sector. In the early stage, AI literacy education often took the form of technology companies working with universities to organize online or offline AI-themed training for faculty and students. Such training reached a broad audience and could provide some introductory knowledge for faculty and students lacking basic AI cognition. However, these training programs had scattered themes, and the education system lacked continuity and systematicity, making it difficult to exert a profound overall impact on improving AI literacy[10]. Subsequently, international organizations and research institutions successively introduced AI competency frameworks for faculty and students, providing macro-level guidance for the design of AI literacy education, such as UNESCO' s AI competency framework. In addition, some educational institutions and libraries have also begun to conduct practical explorations of AI literacy education based on their own capabilities and strengths. For example, many foreign university libraries have carried out AI literacy education through LibGuides[11-12], providing resource guides on topics such as AI tool use and privacy; Japan' s Ministry of Education, Culture, Sports, Science and Technology has launched MDASH to ensure the quality of AI literacy education and promote standardized curriculum development[13]; and in China, a growing number of university libraries are cultivating users' AI thinking through lectures, thematic activities, and other means, while providing targeted AI skills training to enhance their AI application capabilities[14-15].

However, amid the vigorous development of AI literacy education practice, current academic attention is mainly focused on how to construct AI literacy education systems and practical pathways[16-17], while research on evaluation mechanisms for educational activities themselves has not…

timely follow-up. The specific problems are as follows: First, the evaluation content loses focus and lacks a full-cycle perspective. Existing evaluations mostly concentrate on the final assessment of learners' literacy levels, while neglecting the systematic evaluation of the entire process of educational activities. This "black-box" evaluation makes it difficult to identify the root causes of educational problems. Second, there is a lack of a feedback and adjustment mechanism. Evaluation results cannot be used to infer whether deviations occur in the input stage of educational activities or during implementation, making it difficult to form an effective closed loop of diagnosis and feedback. Third, evaluation data are single-sourced and weak in support. At present, evaluation still relies mainly on traditional questionnaires and self-assessment, lacking cross-system, multi-source objective data support; thus, the scientificity and objectivity of evaluation results are difficult to guarantee.

1.3 CIPP Educational Evaluation Model

The CIPP educational evaluation model was proposed in the 1960s by American evaluation scholar D. L. Stufflebeam et al.[18]. It is a systematic evaluation framework widely applied with decision support as its orientation.

CIPP represents the following meanings: C refers to Context Evaluation, that is, analyzing the social environment, policy requirements, learner needs, and existing problems in which the educational activity is situated, so as to ensure that the educational objectives are reasonable, necessary, and targeted; I refers to Input Evaluation, that is, determining what resources should be allocated and what strategies should be adopted to achieve the established goals, thereby ensuring the adequacy of resources and the operability of the educational activity design; P refers to Process Evaluation, that is, focusing on the actual situation in the implementation of education and providing educators with timely feedback through real-time monitoring of teaching projects, teacher–student interaction, and deviations in educational implementation, so as to adjust educational activity strategies and optimize teaching processes, thereby ensuring the quality and effectiveness of educational activities; P refers to Product Evaluation, that is, measuring the extent to which a project has achieved its expected objectives, including short-term learning outcomes and long-term impacts, and providing a basis for the future continuation and expansion of educational activities.

Compared with traditional evaluation models that focus only on final results, the CIPP model places greater emphasis on the diagnostic function of evaluation, long-term development, and service to decision-making. The model has a clear structure and coherent logic, covers the full project cycle, and requires multidimensional data collection. Since its proposal, it has been widely applied in fields such as curriculum reform, educational technology projects, and teacher professional development[19–20]. However, in practice, traditional CIPP evaluation often relies on subjective data such as interviews and questionnaires, resulting in a certain degree of deficiency in objectivity; the multi-source data of university libraries constitute an appropriate supplement. In summary, to address the lack of evaluation of university AI literacy education mentioned above, this study specifically introduces the CIPP evaluation model and uses the multi-source data foundation of university libraries, with the aim of objectively quantifying the AI literacy education evaluation indicator system based on the CIPP model and paving a practical evaluation path of continuous feedback and optimization for AI literacy education in the university context.

2 Multi-Source Data Foundation for the Evaluation of AI Literacy Education in University Libraries

At present, relying on their long-standing foundation in carrying out information literacy and digital literacy education, university libraries have already played a key role in AI literacy education. The teaching practices from information literacy to digital literacy have accumulated valuable teaching experience for university libraries in conducting AI literacy education. Reusing the existing knowledge developed under library literacy education can realize the transfer and upgrading of teaching paradigms, flexibly applying original experience to AI literacy education. For example, in traditional literacy education, university libraries teach learners how to distinguish true from false information, assess the authority of information sources, and identify false information. Such educational content remains indispensable in the AI era and can even be directly transferred to AI literacy education, guiding learners to conduct critical evaluation and bias identification of AI-generated content.

With the popularization and lightweight development of AI technology, university libraries have actively introduced AI-driven tools[21] to improve and transform knowledge service models for teachers and students. As information resource centers and academic service platforms, university libraries are central hubs for information services and knowledge education. In the process of AI literacy education, they have accumulated massive and heterogeneous information, providing a solid data foundation for evaluating AI literacy education at the current stage. The core of this paper is to introduce the CIPP evaluation model to construct an evaluation system driven by multi-source data from university libraries, systematically integrating data scattered across various scenarios of AI literacy education in university libraries and providing objective, precise, and quantifiable data support for the CIPP evaluation model. In view of the CIPP evaluation model' s demand for full-cycle data, this chapter divides the data sources for evaluating AI literacy education in university libraries into five major categories, so as to facilitate data collection in the practical process.

2.1 Library Resource Usage Data

The browsing and downloading history of electronic materials in library systems, AI-themed websites, and records of borrowing related books can directly reflect learners' interactions with AI-related academic content in informal learning scenarios, such as the number of downloads of AI research papers, visits to specific databases, and interactions with AI topics. By analyzing these data, it is possible to reflect teachers' and students' interest in and urgency regarding AI learning, grasp the extent to which they acquire AI knowledge, and fully understand the breadth and depth of their use of AI educational resources provided by the library, thereby indirectly reflecting the evaluation of resource utilization in university library AI literacy education.

2.2 On-Campus AI Tool Usage Data

With the explosive popularity of lightweight large models such as DeepSeek-R1, more and more universities have successively provided locally deployed AI large-model services for teachers and students. Their usage logs are core data for objectively capturing learners' higher-order AI literacy behaviors. They can record the system response speed and service stability records of on-campus AI tools, and preserve learners' frequency and duration of AI tool use, original prompt (Prompt) content, and historical AI output content. This type of data can objectively reflect teachers' and students' actual operation and application capabilities with AI tools, and can analyze the degree of structuring of user instructions and whether they contain questioning vocabulary, for teachers' and students' practical skills and pro—

···provide the most direct basis for assessing prompt-engineering ability, AI-related critical-thinking ability, and potential overreliance and misuse.

2.3 Teaching Service Data

University libraries have long provided information-literacy education services to faculty and students through course training, special lectures, literacy tests, and other means. As an extension of information literacy, AI literacy education in university libraries also follows this model. The data generated in this process are key to quantifying educational activity inputs and evaluating outcomes; they mainly come from the university library' s learning management system (LMS) or from connected educational administration platforms. Specifically, these data include the completeness of AI course syllabi and lecture notes, the proportion of AI case discussions and hands-on practice, participation in AI course training, homework feedback, and assessment scores. Such data can directly reflect the frequency with which faculty and students participate in formal AI literacy education activities and their immediate growth in knowledge and skills, and they also constitute an important basis for measuring the effectiveness of AI literacy education at different stages.

2.4 Library Configuration Data

Library-related configuration is a major strategic basis for evaluating AI literacy education in university libraries, and it is also an important support for diagnosing the quality of resource investment in AI literacy education projects and their adaptability to the contextual environment. It mainly includes: inventories of software and hardware configurations such as the number of library AI workstations and network load capacity; AI-related training records and qualification certificates of teaching librarians; holdings of AI-themed databases and e-books; and the proportion of procurement funds allocated to AI-themed digital resources.

2.5 Subjective Feedback Data

Subjective feedback data are obtained through structured questionnaires, semi-structured interviews, and other forms, and are used to capture learners' subjective perceptions, attitudes, and satisfaction regarding AI literacy education. Such data are difficult to quantify through behavioral data and thus complement the objective behavioral data described above, enabling a multidimensional and comprehensive evaluation of AI literacy education. These data specifically include learners' levels of anxiety about AI technologies and willingness to learn; satisfaction ratings regarding the school' s educational activities and the professionalism of instructors; subjective understanding of AI academic and theoretical norms; and evaluations of their own AI literacy levels.

3 Multi-Source Data Fusion Mechanism for Evaluating AI Literacy Education in University Libraries

To evaluate current AI literacy education activities in university libraries and, with the help of university libraries, achieve a leap in the effectiveness of AI literacy education in higher education, this study proposes conducting fusion analysis and evaluation based on multi-source data obtained during the process of carrying out AI literacy education through university libraries. However, although the five major categories of data sources described above cover all links of AI literacy education, their data types are diverse; they are often scattered throughout the university library knowledge-service process in the form of data silos, and they differ greatly from one another. Accordingly, the multi-source data fusion mechanism proposed in this study mainly consists of three parts: cross-system data association, temporal alignment and data standardization, and feature-engineering transformation oriented to the CIPP model (see Figure 1).

Text in Figure 1: Multi-source data fusion mechanism for evaluating AI literacy education in university libraries

- Data sources:
  - Library resource usage data
  - On-campus AI tool usage data
  - Teaching service data
  - Library configuration data
  - Subjective feedback data
- Encryption
- Access control
- Access control
- Encryption
- Supplement to objective behavioral data
- Anonymous user ID (hash-based desensitization)
- ETL
- Structured data achieve unified semantics and unified storage
- Unstructured data are collected using a unified template
- Synchronous storage of timestamps and time series involving behavioral data
- Feature extraction
- Resource effectiveness → Urgency of real needs → Negative learning behavior → Proficiency and critical thinking
- Dimensions: Context C; Input I; Process P; Product P

Figure 1 Multi-source data fusion mechanism for evaluating AI literacy education in university libraries

3.1 Cross-System Data Association

The core objects of the AI literacy education evaluation constructed in this paper are learners and the educational activities in which they participate. It is therefore necessary to break down the barriers among different business systems within university libraries and establish a learner-centered associated view of all data. In general, library management systems, on-campus AI tool platforms, teaching service systems, and the like are often not in the same system, and there is no unified linkage path among them. Because of the lack of such cross-system mapping, the data are fragmented, making it difficult to support the AI literacy education evaluation proposed in this paper. In on-campus systems, university faculty and students commonly use unique identifiers such as faculty/student IDs and student numbers, which makes it possible to manage data from different sources. Under the premise of protecting faculty and student privacy data and ensuring data ethics, this study uses hash-desensitized anonymous user IDs as the unique identity anchors for different data sources, horizontally integrating the five categories of data sources in a comprehensive manner so as to avoid the direct exposure of sensitive information or

from being reverse-decrypted. At the same time, through encryption and access-control restrictions, the security of data during circulation and invocation is ensured; data are used in a reasonable and compliant manner, thereby promoting the sound development of AI literacy education.

Under the mapping of a unique identity anchor, the system can integrate discrete data from individual dimensions into a relatively complete data network, laying the foundation for feature extraction in AI literacy education activities. For example, linking the "course attendance records" in teaching-service data with the "duration of visits to AI-related topics" in library-resource usage data makes it possible to distinguish between two learner groups: nominal participants and deeply engaged participants. Comparing the "high AI anxiety" found in subjective feedback data with records of "low-frequency use" in on-campus AI-tool usage data can reveal learners' potential separation between knowledge and action.

3.2 Temporal Alignment and Data Standardization

Educational activities have obvious cyclical characteristics, and AI literacy education is no exception. The validity of evaluation data is, to a large extent, constrained by the time window of educational activities. Simple data accumulation may seriously distort evaluation results. For instance, counting AI-tool use during non-teaching periods as part of AI literacy education evaluation would produce vague statistical results that greatly interfere with judgments about the effects of teaching interventions. Therefore, this study introduces a strict temporal-alignment logic into the integration mechanism: timestamps and time series of behavioral data are stored synchronously, so that changes in learners' data indicators over the duration of educational activities can be clearly presented.

AI literacy education requires comprehensive evaluation. The data collected in this paper come from diverse sources, and there are substantial differences among them in format, structure, and semantics; therefore, these data must be systematically integrated and processed. This paper uses ETL (Extract, Transform, Load) tools to process the raw data of university libraries, so as to support later feature-analysis work. In the extraction stage, corresponding interface schemes are designed for data from different sources—for example, retrieving borrowing and access logs from the library management system, obtaining users' course-learning and assignment-grade data from the LMS, and exporting usage logs from AI-tool backends. Obtaining all these data requires connection through specific interface schemes. In the transformation stage, the main problems addressed are inconsistent fields, different encoding methods, and differences in time formats; semantic unification of data from different sources is achieved through data cleaning and standardization. In the loading stage, the processed data are imported into a unified storage platform, forming datasets available for system analysis and invocation. In addition to structured data, semi-structured or unstructured data such as online questionnaires and interview records also need to be collected. Such data carry the subjective attitudes, perceived experiences, and cognitive judgments of teachers and students. They are a powerful supplement and reference for objective behavioral data, and are equally crucial in AI literacy education evaluation. They can be collected through unified templates to maintain field consistency.

3.3 Feature-Engineering Transformation Oriented to the CIPP Model

The ultimate purpose of data fusion is to output quantitative features that can serve AI literacy education evaluation, mapping the cleaned multi-source data into features across four dimensions: context, input, process, and product. In the context and input evaluation dimensions, the focus is on the degree of match between supply and demand. For example, the "resource procurement list" in university library management data can be compared and calculated in stages against the "actual access popularity" in library-resource usage data, so as to diagnose whether current educational-resource allocation is reasonable and precise, and further to observe features of resource effectiveness. Similarly, combining AI anxiety in subjective feedback data with tool-use frequency in on-campus AI-tool usage data can help judge the scientificity of educational-goal setting and construct features reflecting the urgency of real demand.

In the process evaluation dimension, course sign-in data alone can hardly reflect learners' actual learning status comprehensively. To quantify learners' learning depth more accurately, teaching-service data must be integrated and analyzed together with resource-interaction behaviors. For example, the task-completion rate in the teaching-service system can be weighted and fused with course-grade information to present features of AI-knowledge learning efficiency, sensitively capturing negative learning behaviors such as "attendance without engagement," thereby providing early-warning evidence for real-time intervention in the teaching process.

For the product evaluation dimension, the core task is to realize an objective quantitative evaluation of learners' staged improvement in AI literacy. On the one hand, teaching-service data before and after the activity can be used to calculate grade differences, which are transformed into features of knowledge growth. On the other hand, natural-language processing analysis can be conducted on on-campus AI-tool usage data to parse the proportion of structured instructions in prompts and the frequency of questioning vocabulary, thereby obtaining features such as learners' proficiency in AI questioning and critical thinking. Only by combining grade data based on time-series analysis with practical behavioral data can the educational goal of comprehensively improving AI literacy in the age of artificial intelligence be truly achieved.

4 Construction of the AI Literacy Education Evaluation Indicator System for University Libraries and Determination of Weights

4.1 Design of the AI Literacy Education Evaluation Indicator System

At present, research by the academic community and related institutions on AI literacy evaluation systems shows a flourishing trend. Cai Yingchun et al.[6] introduced the KSAVE model and comprehensively evaluated AI literacy from five dimensions: knowledge, skills, attitudes, values, and ethics. Huang Ruhua et al.[22] constructed an AI literacy system from four aspects: artificial-intelligence cognition, artificial-intelligence skills, artificial-intelligence application, and artificial-intelligence ethics. Other scholars, focusing on the core capabilities of GenAI tools, have proposed the subdivided concept of "prompt literacy" from the capability dimension of prompts[23–24], and have used a "pathfinding" model to guide the shaping process of prompt literacy[25]. These evaluation systems all focus on measuring the AI literacy level of individual learners. However, in specific teaching environments—especially in the particular context of university libraries—there is an evident lack of systematic evaluation of the full cycle of AI literacy education activities, making it difficult to address such key questions as whether educational-resource allocation is reasonable and whether the teaching process is effective.

yield valid responses.

As a dynamic and complex systems project, the evaluation of AI literacy education should not be confined to end-point verification of results; rather, it should run through the entire process of educational activities. To this end, this paper introduces the CIPP evaluation model. Taking into account the actual conditions of AI literacy education activities and based on the educational setting of university libraries, it designs a full-cycle evaluation indicator system covering context, input, process, and outcomes. By integrating multi-source data, this system remedies the deficiencies of traditional self-assessment scales and enables an objective and dynamic evaluation of AI literacy education. Based on the CIPP evaluation model, this paper constructs first-level indicators and, accordingly, refines second-level indicators and measurable third-level indicators. It maps multi-source data in university library literacy education to the third-level indicators, thereby forming a complete evaluation system and enabling a three-dimensional view of educational activities, as shown in Table 1.

Table 1 Indicator System for Evaluating AI Literacy Education

First-level indicator	Second-level indicator	Third-level indicator	Corresponding multi-source data and examples
A Context evaluation	A1 Learner needs analysis	A1-1 Baseline level of AI literacy among teachers and students	Teaching service data: pre-instruction diagnostic test scores on AI knowledge among teachers and students, etc.Subjective feedback data: self-evaluation of one' s own AI literacy level
A Context evaluation	A1 Learner needs analysis	A1-2 Interest in and urgency of AI education among teachers and students	Subjective feedback data: self-assessment scores in questionnaires on AI anxiety, training demand, etc.Library resource usage data: downloads of AI research papers, visits to specific databases, and interactions with AI-themed topics, etc.

First-level indicator	Second-level indicator	Third-level indicator	Corresponding multi-source data and examples
B Input evaluation	B1 Quality of resource support	B1-1 Richness of AI educational resources	Library configuration data: holdings of AI-themed databases and e-books, and the proportion of procurement expenditure for AI-themed digital resources, etc.
B Input evaluation	B1 Quality of resource support	B1-2 Availability of on-campus AI tools	On-campus AI tool usage data: system response speed of campus AI tools, records of service stability, etc.
B Input evaluation	B1 Quality of resource support	B1-3 Professional competence of teaching staff	Library configuration data: records of AI-related training for teaching librarians, etc.Teaching service data: teaching evaluation data for instructors in the teaching service system
B Input evaluation	B2 Program design	B2-1 Scientific rigor of curriculum design	Teaching service data: completeness and logical coherence of AI course syllabi and lecture notes, etc.

First-level indicator	Second-level indicator	Third-level indicator	Corresponding multi-source data and examples
B Input evaluation	B2 Program design	B2-2 Diversity of teaching measures	Teaching service data: proportion of AI case discussions and hands-on practice in teaching courses, etc.
C Process evaluation	C1 Attractiveness of teaching activities	C1-1 Actual participation in courses/lectures	Teaching service data: data on completion of AI courses in the LMS, etc.
C Process evaluation	C2 Evaluation of learners' course-completion quality	C2-1 Completion of practical tasks	On-campus AI tool usage data: average number of conversation turns per person with on-campus AI tools, etc.
C Process evaluation	C2 Evaluation of learners' course-completion quality	C2-2 Quality of learning-task completion	Teaching service data: overall AI course grade evaluation, etc.
D Outcome evaluation	D1 Learners' literacy level	D1-1 Technical knowledge	Teaching service data: analysis of gains in knowledge and skills before and after instruction, etc.
D Outcome evaluation	D1 Learners' literacy level	D1-2 Practical skills	On-campus AI tool usage data: growth rate in users' average prompt length before and after education, proportion of structured instructions, etc.

First-level indicator	Second-level indicator	Third-level indicator	Corresponding multi-source data and examples
D Outcome evaluation	D1 Learners' literacy level	D1-3 Critical thinking	On-campus AI tool usage data: number of times AI is questioned during the use of AI tools, etc.
D Outcome evaluation	D1 Learners' literacy level	D1-4 Ethical awareness	Subjective feedback data: questionnaires on attitudes toward AI ethics, etc.
D Outcome evaluation	D2 Satisfaction with teaching activities	D2-1 Satisfaction with courses/instructors	Subjective feedback data: learners' ratings of teaching satisfaction with courses/instructors, etc.
D Outcome evaluation	D3 Long-term impact	D3-1 Consolidation of AI-assisted research/learning behaviors	On-campus AI tool usage data: frequency with which learners use on-campus AI tools

4.2 Weighting of AI Literacy Education Evaluation Indicators

In actual evaluation, AI literacy education evaluation not only requires a clear indicator system; the importance of each dimension under that system and the weights of indicators at all levels are equally important. The AHP method was proposed in the 1970s by the American operations researcher Thomas L. Saaty. It is a decision-making tool that combines qualitative and quantitative analysis. Its core lies in constructing a hierarchical structural model, building pairwise comparison judgment matrices, calculating weight vectors, and conducting consistency tests, ultimately obtaining the relative-importance weights of each indicator. The AHP method is suitable for handling multi-level and complex decision-making problems. It can achieve objective and reasonable weight allocation among different indicators, and can also transform experts' subjective experience into quantitative values. It is well suited to the multidimensional system of AI literacy education evaluation; especially in the context of university libraries, it can effectively integrate expert opinions. Therefore, this paper selects the AHP method to weight the AI literacy education evaluation indicators.

4.2.1 Constructing the Hierarchical Structural Model

Based on the proposed AI literacy education indicator evaluation system, this paper constructs a hierarchical structural model. The target layer is the AI literacy education evaluation indicator system; the criterion layer includes four first-level indicators: background evaluation, input evaluation, process evaluation, and outcome evaluation; the scheme layer is further subdivided into an indicator layer and a factor layer, among which the indicator layer contains eight second-level indicators, and the bottom factor layer consists of sixteen third-level indicators. A questionnaire titled "Weight Allocation for the AI Literacy Education Evaluation Indicator System" was designed using the 1–9 relative scale method. Eighteen experts and scholars in the field of information resource management were invited to participate in the survey, and the experts were asked to conduct pairwise comparisons of the indicators established in this paper and assign values to the relative importance of the indicators. The selected respondents were all experts and scholars familiar with AI literacy education in universities, including 2 university library directors and 16 experts with three or more years of experience in AI literacy education activities.

4.2.2 Calculation of Evaluation Indicator Weights

To determine weights at all levels, it is first necessary to calculate the maximum eigenvalue and the corresponding eigenvector for each judgment matrix. After normalization, local weights are obtained; hierarchical single ranking is then conducted, and the ranking results must pass the consistency test. Taking only the first-level indicators as an example, the judgment matrix is shown in Table 2. The specific calculation process is as follows; the calculation methods for the remaining indicators are the same.

Table 2 Judgment Matrix for the First-Level Indicators of the AI Literacy Indicator Evaluation System

First-level indicator	A Background evaluation	B Input evaluation	C Process evaluation	D Outcome evaluation
A Background evaluation	1	1.555	0.844	0.950
B Input evaluation	0.643	1	0.761	0.923
C Process evaluation	1.185	1.313	1	1.184

First-level indicator	A Background evaluation	B Input evaluation	C Process evaluation	D Outcome evaluation
D Outcome evaluation	1.053	1.084	0.844	1

(1) Hierarchical single ranking

This paper adopts the square-root method commonly used in practical calculation. The steps are as follows:

First, use formula (1) to calculate the product M_i of the elements in each row of the judgment matrix:

$$M_i = \prod_{j=1}^{n} a_{ij} \tag{1}$$

Second, use formula (2) to calculate the fourth root of M_i to obtain $\bar{w}_i$:

$$\bar{w}_i = \sqrt[n]{M_i} \tag{2}$$

Third, use formula (3) to normalize $\bar{w}_i$, obtaining the eigenvector $w_i = [1.057, 0.820, 1.165, 0.991]^{\mathrm{T}}$:

$$w_i = \frac{\bar{w}_i}{\sum_{j=1}^{n} \bar{w}_j} \tag{3}$$

The final w_i obtained is the local weight of the i-th indicator.

(2) Consistency test

After calculating the weight vector, a consistency test must be conducted to examine whether the logic of the expert judgments is consistent. The specific steps are as follows:

First, use formula (4) to calculate the maximum eigenvalue $\lambda_{\max} = 4.018$:

$$\lambda_{\max} = \frac{1}{n} \sum_{i=1}^{n} \frac{(AW)_i}{w_i} \tag{4}$$

Second, use formula (5) to calculate the consistency index $CI = 0.006$:

$$CI = \frac{\lambda_{\max} - n}{n - 1} \tag{5}$$

Third, look up the average random consistency index RI. The RI value can be obtained from a table and is related to the order of the matrix. In this paper, the order of the first-level indicator judgment matrix is 4, and the RI value is 0.8931, as shown in Table 3.

Table 3 Average Random Consistency Index RI Values

Order (n)	1	2	3	4	5	6
RI value	0	0	0.5149	0.8931	1.1185	1.2494

Fourth, using formula (6), calculate the consistency ratio $CR = 0.007$:

$$CR = \frac{CI}{RI} \tag{6}$$

When the CR of the first-level indicator judgment matrix is < 0.1, this indicates that the judgment matrix satisfies acceptable consistency and that the weights of the indicators fall within a reasonable range.

Repeat the above calculation steps to compute the indicator weights at each level, and conduct a consistency test for each judgment matrix. For a multi-level structure, the overall weights of indicators at each level must be calculated through hierarchical overall ranking: the composite weight of an indicator at each level is the product of its local weight and the overall weight of the indicator at the level immediately above it. This ultimately yields the weight-coefficient table. As shown in Table 4, all obtained weight-coefficient results passed the consistency test.

Table 4 Weights of the Evaluation Indicator System for AI Literacy Education

First-level indicator	Weight	Second-level indicator	Weight	Third-level indicator	Weight
A Background evaluation	0.262 1	A1 Learner needs analysis	0.262 1	A1-1 Baseline level of AI literacy among teachers and students	0.156 3

First-level indicator	Weight	Second-level indicator	Weight	Third-level indicator	Weight
A Background evaluation	0.262 1	A1 Learner needs analysis	0.262 1	A1-2 Interest in and urgency of AI education among teachers and students	0.105 8
B Input evaluation	0.203 3	B1 Quality of resource support	0.134 8	B1-1 Richness of AI education resources	0.073 0
B Input evaluation	0.203 3	B1 Quality of resource support	0.134 8	B1-2 Availability of on-campus AI tools	0.024 6
B Input evaluation	0.203 3	B1 Quality of resource support	0.134 8	B1-3 Professional competence of instructors	0.037 2
B Input evaluation	0.203 3	B2 Program design	0.068 5	B2-1 Scientific rigor of curriculum design	0.042 2
B Input evaluation	0.203 3	B2 Program design	0.068 5	B2-2 Diversity of teaching measures	0.026 3

First-level indicator	Weight	Second-level indicator	Weight	Third-level indicator	Weight
C Process evaluation	0.288 9	C1 Attractiveness of teaching activities	0.159 3	C1-1 Actual participation in courses/lectures	0.159 3
C Process evaluation	0.288 9	C2 Evaluation of the quality of learners' course completion	0.129 6	C2-1 Completion of practical tasks	0.056 8
C Process evaluation	0.288 9	C2 Evaluation of the quality of learners' course completion	0.129 6	C2-2 Quality of learning-task completion	0.072 8
D Outcome evaluation	0.245 7	D1 Learner literacy level	0.084 9	D1-1 Technical knowledge	0.019 6
D Outcome evaluation	0.245 7	D1 Learner literacy level	0.084 9	D1-2 Practical skills	0.027 6
D Outcome evaluation	0.245 7	D1 Learner literacy level	0.084 9	D1-3 Critical thinking	0.022 8
D Outcome evaluation	0.245 7	D1 Learner literacy level	0.084 9	D1-4 Ethical awareness	0.014 9

First-level indicator	Weight	Second-level indicator	Weight	Third-level indicator	Weight
D Outcome evaluation	0.245 7	D2 Satisfaction with teaching activities	0.099 1	D2-1 Satisfaction with the course/lecturer	0.099 1
D Outcome evaluation	0.245 7	D3 Long-term impact	0.061 7	D3-1 Consolidation of AI-assisted research/learning behaviors	0.061 7

5 Practical Paths for Evaluating AI Literacy Education in University Libraries

The significance of educational evaluation lies not only in presenting results, but also in systematically analyzing whether educational activities operate well across different dimensions. The introduction of multi-source data provides educational diagnosis with evidence of greater depth and breadth than before. Evaluation data are no longer confined to learners' self-reported feedback; instead, they shift toward integrated analysis linking behavioral data, process data, and outcome data. At the same time, the pursuit of improvement is also the original intention behind the CIPP evaluation model. Therefore, based on the multi-source data fusion mechanism and evaluation indicator system constructed above, this section conducts an effective diagnosis of AI literacy education in university libraries, realizes visualization and dynamic feedback of evaluation results for AI literacy education activities, and, on this basis, proposes improvement strategies for the continuous optimization of AI literacy education.

5.1 Diagnosis of Evaluation Results for AI Literacy Education

Judging the overall effectiveness and potential problems of university libraries in AI literacy education is a prerequisite in the implementation of AI literacy education in universities. From the dimension of background evaluation, cross-analysis of learners' pre-education baseline level, demand preferences, and usage behavior can reveal the degree of fit between current educational goals and learners' actual conditions. For example, when questionnaire feedback shows that learners have strong anxiety about AI technology, while their use of AI resources remains low over the long term, this indicates that there is a certain deviation between the educational goals set at the initial stage and students' real needs. This may be because the university library has failed to effectively meet learners' needs in terms of course guidance or resource presentation. The input-evaluation dimension relies on indicators such as resource support, instructor competence, and curriculum design for comprehensive judgment. Through the integrated analysis of multiple indicators, it becomes possible to more clearly determine whether there is a significant gap between the university library' s resource allocation and learners' access preferences—for example, some AI-themed databases may have been introduced at high cost but are rarely used in any substantive way.

The process-evaluation dimension focuses on the practical state of teaching activities, such as identifying whether learners exhibit "surface participation" by integrating and analyzing course participation, task completion rates, and resource-interaction behaviors. In addition, if the frequency of AI tool use during a course shows a significant discontinuity, this may mean that the teaching rhythm or course-task design has failed to generate sustained appeal. The outcome-evaluation dimension starts from the perspective of literacy improvement and specifically observes the changes, before and after AI literacy education activities, in the four core elements of AI literacy defined earlier—technical knowledge, practical skills, critical thinking, and ethical awareness—thereby providing an effective basis for university libraries to judge the stage-specific effectiveness of AI literacy education.

5.2 Visualization Analysis of Evaluation Results and Dynamic Feedback

In the process of promoting AI literacy education, university libraries often face problems such as multiple teaching teams, complex task types, and dispersed data. Ordinary text-based evaluation reports are difficult to use in supporting convenient decision-making. Only visualized evaluation results can more clearly present the current pain points and development trends of educational activities, providing librarians and teachers with more actionable bases for decision-making. Taking the four dimensions of the CIPP evaluation model as the main thread, radar charts can be used to display the overall evaluation results for context, input, process, and product, enabling university library managers to accurately identify weak links in the AI literacy education system. During the implementation of teaching activities, data such as time-series resource access volume, course participation, and frequency of use of campus AI tools can be used to intuitively analyze whether the teaching rhythm and learner activity remain consistent. Visualizing behavioral features such as the structured proportion of prompts and the frequency of questioning expressions provides more concrete references for teaching managers to examine learning outcomes.

Promoting the establishment of a dynamic feedback mechanism is another value of visual analysis. Based on visualized data results, university libraries can provide phased feedback on the implementation process of AI literacy education activities, promptly conveying learners' stage-by-stage performance to teaching managers so that the structure of teaching content can be adjusted, the teaching pace controlled, or teaching resources supplemented. For example, when the visualized structure of behavioral data shows a cliff-like decline in the depth and frequency of students' use of AI tools during a certain period, relevant managers can judge accordingly whether practical guidance needs to be strengthened in the next teaching segment; when test results on ethical awareness remain at a relatively low level for a long time, the proportion of ethics-related topics in teaching activities should be appropriately increased.

The construction of a dynamic feedback mechanism helps break through the limitations of one-off AI literacy education evaluation, lays out a traceable path for continuous improvement, and forms a dynamic, long-term AI literacy education system.

5.3 Closed-Loop Optimization and Continuous Improvement Strategies

To truly improve AI literacy education, it is not enough to remain at the level of superficial reflection after evaluation; instead, the evaluation process must operate in a closed loop and support continuous improvement. Through evaluation diagnosis and analysis based on continuous multi-source data input, the structure of AI literacy education can be promoted to evolve continuously, enabling university libraries to gradually transform from traditional activity organizers into long-term operators of AI literacy enhancement. Based on the multi-source data fusion mechanism and evaluation indicator system proposed in this paper, a cyclical chain covering the entire process of AI literacy education can be established, thereby improving university libraries' ability to correct deviations in the AI literacy education process.

Optimization in the context-evaluation dimension lies in dynamically revising objectives and accurately identifying learners' cyclical needs. When there is a serious divergence between learner behavior and subjective feedback results, university libraries can reposition educational objectives so that they better align with learners' current actual ability status. In terms of the input-evaluation dimension, university libraries can re-examine resource allocation and teaching-staff conditions according to the evaluation results—for example, by increasing AI resources that are highly used but insufficiently supplied, or by strengthening librarians' AI competence training in a targeted manner based on feedback, so as to improve the matching degree of teaching supply. Improvements in the process-evaluation dimension are mainly reflected in the upgrading of teaching methods. When a learning gap or insufficient engagement appears in a certain segment of AI literacy education, university libraries can promote learners' depth of participation by adjusting the structure of practical tasks and optimizing the course rhythm. In the product-evaluation dimension, weaker ability indicators serve as key directions for improving future AI literacy education courses. For example, the widely existing problem of insufficient critical thinking will prompt university libraries to strengthen analysis of, or reflection on, relevant cases during AI literacy education activities.

Only by forming a long-term and stable closed-loop mechanism for AI literacy education operations can university libraries realize the normalization of AI literacy education activities. On this basis, university libraries can establish fixed evaluation cycles and data governance mechanisms, gradually constructing a development path with distinctive AI literacy education characteristics suited to their own institutional realities through continuous data accumulation and longitudinal comparison. This will promote closer coordination and cooperation between university libraries and teaching departments, and help university libraries become learning hubs for future universities.

6 Conclusion

In the era of artificial intelligence, AI literacy education has become one of the indispensable core educational themes. Although carrying out AI literacy education is certainly important, educational evaluation is equally indispensable; only scientific and reasonable evaluation can bring about positive changes in the educational process. This paper systematically describes the paradigm shift of AI literacy, summarizes and explains its specific connotations in the context of higher education, and analyzes the current status and challenges of AI literacy education. Relying on the attributes of university libraries as learning centers and information hubs, it forms a multi-source data foundation in order to construct a more complete multi-source data fusion mechanism for evaluating AI literacy education in university libraries. At the same time, the CIPP evaluation model is introduced to design multidimensional evaluation indicators that can be objectively quantified by actual data, and the AHP method is used to assign weights to indicators at all levels, with the aim of providing a referential evaluation indicator system for AI literacy education assessment in university libraries and further improving the quality of AI literacy education. This paper also has certain limitations: the evaluation system proposed in the article remains at the theoretical stage and has not yet been practically developed, so its effectiveness and feasibility require further verification. In the future, it will be possible to explore how to realize shared data and practical cooperation among different university libraries, jointly develop standardized AI literacy education evaluation tools, and help AI literacy education advance in a more comprehensive and extensive direction.

References

[1] Zhang Zheng, Liu Chenxu. Large-model hallucination: Cognitive risks and possibilities for co-governance in human–machine communication[J]. *Journal of Soochow University (Philosophy and Social Sciences Edition)*, 2024(5): 171–180.

[2] Bu Miao. Practice and pathways of AI literacy education in university libraries[J]. *Library Work and Study*, 2025(2): 89–96, 112.

[3] Zurkowski P G. The information services environment relationships and priorities. Related Paper No. 5[R]. Washington, D.C.: Na-

tional Commission on Libraries and Information Science, 1974:2.

[4] American Library Association. Framework for information literacy for higher education [EB/OL]. [2025-09-03]. http://www.ala.org/acrl/standards/ilframework.

[5] UNESCO. A global framework of reference on digital literacy skills for indicator 4.4.2 [EB/OL]. [2025-09-03]. https://unesdoc.unesco.org/ark:/48223/pf0000265403.

[6] Cai Yingchun, Zhang Jingbei, Yu Chenlin, et al. Artificial intelligence literacy in the era of digital intelligence: connotations, framework, and implementation paths [J]. *Journal of Library Science in China*, 2024(4): 71–84.

[7] Xin Bingxin, Ye Lan, Hu Yansong. A survey of artificial intelligence literacy education in foreign university libraries: an analysis based on artificial intelligence LibGuides [J]. *Journal of Academic Libraries*, 2024(4): 58–68.

[8] Official Journal of the European Union. Regulation (EU) 2024/1689 of the European parliament and of the council of 13 June 2024 [EB/OL]. [2025-09-03]. https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng.

[9] Chinese Social Sciences Net. Strengthening guidance on the standardized use of generative AI tools [EB/OL]. [2025-09-03]. https://www.cssn.cn/skgz/bwyc/202505/t20250506_5872424.sht

[10] Zhang Miao. A survey of AI literacy education practices in university libraries in China [J]. *Library and Information Service*, 2025(18): 98–108.

[11] Du Xiuxiu, Xu Bowen, Chu Jiewang. Research on generative artificial intelligence resource navigation in libraries of top research universities in the United States [J]. *Library Journal*, 2025(3): 68–81.

[12] Li Teng. Practices and implications of artificial intelligence literacy education in university libraries in the United Kingdom and the United States [J]. *Library Work and Study*, 2025(3): 49–55, 86.

[13] Dai Xinwei, Li Feng. Design, practice, and implications of artificial intelligence literacy education in Japanese universities under the MDASH certification background [J]. *Journal of Library and Information Science in Agriculture*, 2025(5): 86–101.

[14] Chen Fangbin. Current status and improvement strategies of AI literacy education in Chinese university libraries: based on a survey of 19 research university libraries [J]. *Library Work and Study*, 2025(3): 75–86.

[15] Xu Ying, Gao Tian, Cheng Xiao. A case study of AI literacy education in university libraries based on content analysis [J]. *Library World*, 2025(1): 5–10, 32.

[16] Zhang Jingbei, Yu Chenlin, Cai Yingchun, et al. Research on the construction of an AI literacy education content system for research scenarios [J/OL]. *Library Journal*, 1-17 [2025-09-30]. https://link.cnki.net/urlid/31.1108.G2.20250218.1330.002.

[17] Zhang Yi. Exploring pathways for AI literacy education in university libraries from a global perspective [J]. *Journal of Academic Libraries*, 2025(5): 86–97.

[18] Stufflebeam D L, Shinkfield A J. *Systematic evaluation: a self-instructional guide to theory and practice* [M]. Dordrecht: Kluwer-Nijhoff Publishing, 1985.

[19] Xiao Peng, Ge Zhoufeng, Hao Xue. Exploration of the OBE+CIPP classroom evaluation model [J]. *Research in Higher Education of Engineering*, 2021(6): 176–182.

[20] Li Tong, Zou Xiaozhu. Research on the evaluation of collaborative information literacy education in universities based on the CIPP model: taking a course at Nanjing University of Aeronautics and Astronautics as an example [J]. *Information and Documentation Services*, 2016(2): 108–112.

[21] Guo Yajun, Xu Yuanqian, Liang Yanli, et al. From ChatGPT to DeepSeek: the impact of generative artificial intelligence iteration on libraries [J]. *Library Tribune*, 2025(7): 140–149.

[22] Huang Ruhua, Shi Leyi, Wu Yingqiang, et al. Construction of a content framework for artificial intelligence literacy education in China from a global perspective [J]. *Documentation, Information & Knowledge*, 2024(3): 27–37.

[23] Zhao Yuxiang, Jing Litian, Song Shijie, et al. Prompt literacy empowered by AIGC: reconstructing human–AI interaction capabilities in the era of generative AI [J]. *Information and Documentation Services*, 2025(3): 14–25.

[24] Zhang Guixiang, Jia Junzhi. Research on the cultivation of prompt literacy in the era of generative AI [J]. *Journal of Academic Libraries*, 2024(6): 63–71.

[25] Cui Yuhong, Qu Wenlan, Wang Sa, et al. "Pathfinding" new-quality education: the theoretical framework, practical pathways, and future prospects of prompt literacy [J]. *Journal of Academic Libraries*, 2025(2): 39–49.

Research on Multi-source Data Fusion-Based AI Literacy Education Evaluation in University Libraries

Chu Jiewang1 **Zhou Keyan**1 **Niu Yue**2

1. School of Management of Anhui University, Hefei, 230601

2. Northwestern Polytechnical University Library, Xi' an, 710072

Abstract With the rapid advancement of generative AI (GenAI), AI literacy has become one of the core competencies to be cultivated in higher education. As key learning hubs, university libraries are actively conducting AI literacy education. However, the current university community has not yet prioritized the evaluation mechanism for AI literacy education, resulting in insufficient maturity in effectively assessing such educational initiatives within university libraries. From the perspective of academic libraries, this study systematically reviews the evolution and connotation of AI literacy education, analyzes existing problems, and clarifies the unique advantages of university libraries in delivering AI literacy education. The CIPP evaluation model is introduced to construct a multi-source data-driven evaluation system. By effectively mapping multi-source data from university libraries onto the four dimensions of the CIPP evaluation model, the study proposes a comprehensive and quantifiable indicator system for AI literacy education evaluation, enabling objective and dynamic assessment of AI literacy education activities in university libraries. Furthermore, the Analytic Hierarchy Process (AHP) method is employed to determine the corresponding weights of indicators at all levels, providing theoretical support and practical pathways for university libraries to continuously optimize their AI literacy education services.

Keywords AI literacy; AI literacy education evaluation; University libraries; Evaluation system; CIPP evaluation model