Asia Pacific Science Press

Asia Pacific Economic and Management Review
ISSN(O):3005-9275
ISSN(P):3005-9267
Vol. 1 No.1 (2024)

Application of Internet User Behavior Data in Library Big Data Analysis

Fan Yang*

National Library of China, Beijing, Beijing 100081, China

Corresponding author: Fan Yang

Copyright: 2024 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY-NC 4.0), permitting distribution and reproduction in any medium, provided the original author and source are credited, and explicitly prohibiting its use for commercial purposes.

Abstract: This paper investigates the current state of data utilization in libraries, thoroughly analyzing the characteristics, collection methods, and processing techniques of internet user behavior data. In response to the service demands of libraries, it proposes application models such as user profiling, personalized recommendation services, and optimized collection decision-making. Furthermore, the paper offers suggestions for building a data governance system and an open knowledge service platform, aiming to enhance libraries' data-driven service capabilities. These strategies align with the central government's directive to "promote innovation in big data technology and industry development, build a digital economy driven by data, use big data to modernize national governance, improve public welfare, and safeguard national data security" [1].

Keywords: Big Data Management; Intelligent Application; Library; Management Systems; User Behavior Data; Data Analysis
Published: Feb 15, 2024
DOI: https://doi.org/10.62177/apemr.v1i1.260

1.Introduction

With the rapid advancement and widespread application of information technology, big data has emerged as a critical and strategic resource underpinning the sustainable development of modern society and the broader economy. As essential institutions for the provision of public cultural services and the dissemination of knowledge, libraries are increasingly being called upon to respond to the evolving digital landscape. One of the most pressing and complex challenges they currently face is how to effectively harness and integrate data derived from internet user behavior in order to enhance the scope, precision, and adaptability of their service delivery systems.

At present, there exist notable deficiencies in the data management practices of many libraries, particularly in the domain of user behavior analytics. Most libraries lack a systematic and comprehensive framework for collecting, analyzing, and utilizing behavioral data, which in turn constrains their ability to offer personalized and targeted services that meet the dynamic needs of their users. This gap significantly hampers efforts to modernize service models and respond proactively to shifts in user expectations and information consumption patterns.

In this context, conducting research on how to effectively consolidate and apply internet user behavior data within library operations becomes not only timely but also imperative. Such efforts aim to construct scientific, data-informed service models that can guide future innovations in library management. This line of inquiry carries substantial theoretical implications for the development of data-driven decision-making mechanisms, while also offering practical value for enhancing the relevance, efficiency, and user-centeredness of library services in the era of digital transformation.

2.Scoring Efficiency

2.1 Progress in Foreign Research

Academic and practical exploration into the application of big data in libraries began relatively early in Western countries. Leading institutions such as the Harvard University Library and the Yale University Library have taken the initiative to integrate user behavior data analytics into their operational frameworks. By analyzing patterns in user behavior—including search queries, resource access frequency, and navigation paths—these libraries have achieved notable advancements in the precision of resource allocation and the personalization of user services, thereby significantly enhancing overall user satisfaction [2].

The focal points of these studies and implementations have revolved around three main domains: data governance, optimization of resource configuration, and intelligent recommendation services. Over time, these institutions have developed relatively mature models of data governance, which not only support internal decision-making but also facilitate responsive and user-centered service innovation.

Specifically, the Harvard University Library launched an early-stage reader behavior analytics project grounded in big data methodologies. By collecting and analyzing extensive data on users' visit logs, keyword searches, and resource download behaviors, the library was able to accurately identify patterns of interest and user preferences. This insight directly informed decisions regarding collection development, spatial arrangement of resources, and acquisition strategies, leading to more data-driven and user-aligned resource planning [3].

The Yale University Library has similarly leveraged advanced data mining techniques and machine learning algorithms to extract actionable usage patterns from massive volumes of user behavior data. Based on these insights, the library developed a personalized resource recommendation system, which not only improved the efficiency of information retrieval but also substantially increased user satisfaction and engagement with library collections.

In addition, the Bodleian Libraries at the University of Oxford have adopted big data analytics to support the development of an intelligent book recommendation platform. By analyzing historical borrowing records and online interaction behaviors, the system is capable of making accurate predictions regarding users' informational needs and preferences. This has enabled more intelligent resource recommendation and configuration strategies, contributing to enhanced service precision and operational efficiency.

The National Library of Australia has employed a different but equally effective approach. By integrating data from social media platforms and the library's official website, and applying natural language processing (NLP) techniques, the library has been able to extract latent user demands and interest features from unstructured data sources. This has substantially improved the effectiveness of its digital collections and enhanced the user experience in a measurable way.

Moreover, a wide range of public libraries across Europe and North America have increasingly incorporated user behavior analytics into their routine service optimization processes. Through regular data mining and behavioral analysis, these institutions are able to detect evolving patterns in user needs. The results inform dynamic adjustments in collection development and procurement policies, ultimately leading to marked improvements in community user satisfaction [4].

2.2 Current Research Status and Existing Problems in China

In contrast to the relatively mature big data practices observed in libraries in Western countries, the development of big data applications in Chinese libraries remains predominantly at the stage of theoretical exploration, with practical implementation lagging significantly behind. Current efforts are hindered by a range of structural and technological barriers, including difficulties in data sharing, the widespread existence of data silos, and outdated data collection technologies. As a result, an effective data-driven service system has yet to take shape in most domestic libraries[5].

A closer examination of the situation reveals that research on big data in Chinese libraries is mainly focused on conceptual frameworks and preliminary technological trials, while systematic and replicable case studies remain scarce. On one hand, the issue of data silos is particularly severe. Information systems in many libraries continue to operate independently, with little to no integration between platforms. These systems often lack unified data standards and effective mechanisms for interoperability and data exchange, leading to significant data fragmentation. Such limitations seriously hinder the potential

for comprehensive cross-platform analysis and application, thus undermining the utility of data in enhancing library services.

On the other hand, the technological infrastructure supporting data collection in domestic libraries remains underdeveloped. Many institutions have not yet established robust mechanisms for collecting and analyzing user behavior data—such as search histories, browsing patterns, and online social interactions. Consequently, the immense potential of internet user behavior data for driving service innovation and personalization has not been fully realized [6].

Moreover, there is an evident absence of a comprehensive data governance framework and supporting technical platforms across Chinese libraries. Key challenges include insufficient capabilities in data management and governance, low levels of data standardization, and inadequate development of data security systems. These weaknesses collectively constrain the effectiveness of data utilization and impede the advancement of innovative, data-driven services.

Given these circumstances, there is an urgent need to accelerate the construction of a robust and unified data governance system. This includes the formulation of standardized data protocols, the establishment of effective sharing mechanisms, and the adoption of advanced data collection and analytical technologies. By addressing these foundational issues, libraries in China can significantly enhance their capacity to provide responsive, intelligent, and user-centered services in the era of big data [7].

3.Analysis of Library Data Types and Current Application Status

3.1 Analysis of Library Data Types

Library data is characterized by its diversity and complexity, encompassing a wide range of data types including structured, unstructured, and semi-structured data. Structured data—such as bibliographic records, user borrowing histories, and electronic resource access logs—typically exhibits a well-defined organizational schema and is easily searchable. These features make structured data highly effective for supporting routine library operations, including catalog searching, circulation analysis, collection development, and user behavior profiling.

In contrast, unstructured data refers to information that lacks a fixed format, including images, audio files, video recordings, and scanned documents. Due to its varied forms and massive volume, unstructured data poses significant challenges in terms of storage, processing, and analysis. Nevertheless, it plays a vital role in the development of digital libraries, particularly in the management and utilization of multimedia resources. Tasks such as content indexing, semantic tagging, and media retrieval require advanced data processing capabilities to unlock the full value of these assets.

Between these two extremes lies semi-structured data, which includes formats such as webpages, HTML documents, XML files, and emails. These data types contain structural elements but do not conform to rigid relational models. In the context of libraries, semi-structured data is particularly valuable for tasks related to resource integration, user interaction analysis, and the enhancement of web-based services and communication platforms.

The growing variety and rapid expansion of these heterogeneous data types bring about new challenges for data governance, management, and effective utilization. Traditional data handling methods are increasingly inadequate in addressing the complexity, scalability, and performance demands posed by such data. Consequently, libraries must prioritize the enhancement of their data processing technologies, adopt advanced analytics frameworks, and strengthen their data governance capabilities in order to achieve efficient, secure, and value-driven data use in a rapidly evolving information environment.

3.2 Analysis of the Current Data Status and Problems of the National Library

The National Library of China operates multiple independent business management systems, including ALEPH, the Wenjin Search System, and others, which are responsible for various aspects of library operations. These systems generate substantial volumes of operational data, such as borrowing records, bibliographic and collection data, and reader profile information. However, due to the independent operation of each system, the overall level of data integration remains low, and data sharing across platforms is significantly limited, resulting in a highly fragmented data environment.

One of the critical challenges lies in the lack of standardized interfaces and data exchange protocols between systems. This absence of interoperability hinders seamless data flow and prevents the realization of cross-platform data integration, thereby limiting the potential for comprehensive analytics and coordinated resource management. As a result, valuable insights that

could be derived from correlating data across systems remain untapped.

Furthermore, the absence of a unified data governance framework has led to inconsistent data standards and uneven data quality. Discrepancies in data formats, naming conventions, and update cycles further complicate efforts to utilize the data effectively. The inefficiencies stemming from poor data quality and lack of standardization not only compromise the accuracy of analytical outcomes but also constrain the library's ability to make informed decisions and deliver high-quality, personalized services to its users.

Given these challenges, there is an urgent need for the National Library to establish a comprehensive and unified data governance system. This system should aim to standardize data formats, enforce quality control mechanisms, and create efficient channels for data exchange between systems. Promoting system-wide data fusion and intelligent utilization will be essential for unlocking the full potential of operational data, thereby enhancing the library's overall service capabilities, resource management efficiency, and strategic decision-making capacity in the digital age.

3.3 Analysis and Application of Tag Data

Tag data represents a crucial type of user-centered information that reflects individual interests, behaviors, and preferences. In the context of libraries, tag data can be broadly categorized into three types: static tags, such as age, occupation, and other demographic characteristics; dynamic tags, which capture evolving patterns of user interests and behaviors over time; and predictive tags, which aim to forecast future directions of user interest based on historical data and behavioral trends (Li, 2016). These tag types together serve as a foundational component in the development of personalized services and intelligent resource recommendation systems in modern library environments.

The application potential of tag data in libraries is considerable. By enabling fine-grained user profiling and facilitating targeted content delivery, tag-based systems can significantly enhance user satisfaction, engagement, and service relevance. However, despite its importance, the actual utilization of tag data in library practice remains limited by several critical issues related to data quality and system design.

First, data freshness remains a major challenge. Tag information is often not updated in a timely manner, which results in outdated or inaccurate user profiles. Second, the lack of standardized tag formats and classification rules across systems leads to inconsistency in how user attributes are labeled and interpreted. This lack of standardization not only impedes data interoperability but also undermines the reliability of personalized services across platforms. Third, the granularity and precision of tag content are frequently insufficient. Tags may be too generic or irrelevant to capture the nuanced preferences of users, thereby limiting the effectiveness of recommendation algorithms and weakening the accuracy of user portraits.

These limitations collectively restrict the development of data-driven service capabilities within libraries. To address these challenges, it is imperative to invest in the enhancement of tag data quality by increasing the frequency of updates, establishing uniform tagging standards, and improving semantic accuracy and contextual relevance. Only by advancing these aspects can libraries fully realize the value of tag data in enabling responsive, intelligent, and user-tailored services in the era of big data.

4.Characteristics and Analysis of Internet User Behavior Data

4.1 Dimensional Analysis of User Behavior DataAnalysis and Application of Tag Data

Internet user behavior data encompasses a wide array of dimensions that collectively form the analytical foundation for user modeling and personalized service delivery. These dimensions are typically categorized into three core types: static attributes, interest-based dimensions, and social interaction dimensions. Together, they offer a multi-faceted representation of user behavior and preferences, serving as the basis for advanced data-driven decision-making in digital information environments such as libraries.

Static attributes refer to foundational demographic information, including variables such as gender, age, geographic location, educational background, and occupational category. These data elements form the structural core of a basic user profile, offering initial insights into user segmentation and group characteristics. While relatively stable, static attributes provide essential context for tailoring services to broad user categories and initiating preliminary filtering in recommendation systems.

Interest-based dimensions, on the other hand, capture the dynamic and personalized behavioral patterns of users. These

include data on search queries, browsing habits, and historical purchasing behavior. By analyzing these elements, institutions can gain nuanced insights into individual users' actual preferences and latent needs. For instance, frequent search keywords and their recurrence patterns reveal the thematic interests of users, while metrics such as browsing duration and visit frequency offer clues about the depth of interest and engagement intensity. Additionally, purchasing history—when available—can offer concrete evidence of user intent and behavior, thus allowing for the construction of highly targeted service or content recommendations.

Social interaction data constitutes the third dimension and reflects the user's behavior within online communities and networks. This includes comments, likes, shares, reposts, and other interactive actions across social media platforms or within digital content environments. These social signals are valuable not only for understanding a user's attitude toward specific resources but also for analyzing interpersonal relationships, community influence, and content diffusion pathways. Using methods such as social network analysis (SNA), libraries and information service providers can explore the collective behavior of user groups, track the propagation of interests, and even identify opinion leaders within user communities.

Among these, the analysis of interest-based data is particularly crucial for achieving accurate user profiling and service personalization. For example, frequent analysis of user search terms and their occurrence frequency enables the identification of core interest themes. By further examining browsing duration, click-through behavior, and page revisit patterns, it becomes possible to infer interest intensity and preference orientation. Similarly, analysis of purchase records can yield rich insights into users' material needs and consumption behaviors, leading to more refined and predictive models of user demand.

In the realm of social data, interactive behavior on digital platforms adds another dimension to user analysis. Comments, reposts, and likes not only indicate users' preferences but also reflect their engagement levels and influence within a social network. By mapping out these behaviors through social network analytics, organizations can uncover the dynamics of interest diffusion, detect clusters of shared preferences, and design community-responsive service strategies.

In summary, the integration and analysis of static, interest-based, and social data dimensions are essential for constructing comprehensive user portraits. Such multidimensional profiling allows institutions to move beyond one-size-fits-all approaches and toward context-aware, personalized service ecosystems, enhancing user satisfaction, engagement, and service effectiveness in a data-rich digital age.

4.2 The Theory and Application of Folksonomy

Folksonomy reflects the characteristics of user-initiated tagging, with Douban's social tagging system being a typical application example. Through spontaneous user tagging and evaluation, accurate analysis of user interests and needs can be achieved, which is of significant reference value for the resource construction and service optimization of libraries. Specifically, folksonomy emphasizes user-initiated tagging, allowing users to flexibly express their needs and preferences through free and informal tagging forms [9]. Unlike traditional professional classification methods, folksonomy is more timely and precise in reflecting users' real needs, capable of quickly capturing trends in user interests.

For instance, Douban's successful experience with its social tagging system is worth learning from libraries. Users tag resources autonomously, with tag content covering multidimensional information such as resource characteristics, personal feelings, evaluations, and points of need [10]. These tags not only help users locate the resources they need more quickly but also provide valuable user interest data for libraries. Libraries can conduct data mining and analysis based on user-initiated tagging systems, accurately grasp the trends in user interests, optimize resource acquisition and service strategies, and enhance the personalization of library services [11].

Moreover, by analyzing the relevance and trend changes of user tags, libraries can promptly identify resource hotspots and user needs trends, making precise resource recommendations and service arrangements. This folksonomy based on user-initiated tags effectively compensates for the shortcomings of traditional professional classification methods in personalized services, greatly enhancing the precision and applicability of library resource construction and services.

5.Internet User Behavior Data Collection and Processing Technology

5.1 Principles and Applications of Web Crawler Technology

Web crawler technology is an automated tool for internet data collection, which achieves rapid and efficient data acquisition

through specific programs. This article focuses on the advantages and application processes of focused crawlers in data collection within specific domains.

5.2 Practical Analysis of Data Collection Technology Examples

The Octopus Data Collection System is a powerful web data collection tool capable of efficiently scraping data from specific websites and web page structures, and is easy to operate. Taking the Octopus Data Collection System as an example, the strategies and specific operational steps for collecting internet user behavior data include:

(1) Creating a collection task: First, the user creates a new task in the Octopus system and inputs the URL of the target website.

(2) Setting collection rules: For instance, when collecting user review data from a book e-commerce platform (such as Dangdang), the user needs to select and set the target data fields through the system's built-in visual interface, including user name, review content, review time, rating, and other specific details.

(3) Running the task for automatic collection: After setting the rules, the collection task is initiated, and the Octopus system automatically visits the website pages, extracts user review data according to the preset rules, and automatically flips through multiple pages to capture data.

(4) Data export and processing/cleaning: After collection, the user exports the data into Excel, CSV, or database formats. Based on the exported raw data, the data quality is further improved through data cleaning steps, such as removing duplicate data, cleaning abnormal data, standardizing time formats, and rating criteria.

6.Construction of Library Application Models Based on Internet User Behavior Data

6.1 Construction of Reader and Resource Profiles

Reader profiles are constructed by analyzing internet behavior data of readers, such as browsing history, borrowing history, search keywords, and purchase records, to create a comprehensive and precise user demand model. For instance, by analyzing the frequency of search keywords, search preferences, and the fields of resources that users are interested in, the focus of user interests can be determined. Additionally, by examining the online reading duration and frequency, the intensity of interest in specific content can be assessed. Based on social dimension data, such as user comments, ratings, and resource sharing interactions, user emotional tendencies and social influence can be effectively identified, thereby further enriching the content of reader profiles [12].

Resource profiles are constructed by analyzing the tagged information, user evaluations, and access records of library resources to establish detailed feature tags for the resources. For example, for a specific book resource, combining reader comments and tag annotations can clarify its content characteristics and suitable reader demographics, achieving precise resource description and classification.

It is also noteworthy that during the actual construction process, a unified data standard and model specification should be established to ensure the precision and consistency of reader and resource profiles. Moreover, continuous updates to the profiles should be emphasized to maintain their real-time relevance and precision.

6.2 Personalized Recommendation Service System

The personalized recommendation service system is based on reader and resource profiles, achieving precise push services through intelligent matching algorithms. On one hand, it leverages the user interest preference data provided by reader profiles to accurately select resource content that highly matches user needs; on the other hand, it precisely identifies the user groups that resources are suitable for through the detailed descriptions of resource profiles, realizing a two-way precise match between resources and users.

For example, by analyzing reader historical browsing and borrowing records data, combined with the content attributes and tag features of resources, the intelligent recommendation system can automatically push books and electronic resources that align with reader preferences, significantly improving the efficiency of users obtaining information. At the same time, by continuously analyzing real-time dynamic user behavior data, the recommendation algorithm is constantly optimized, promptly capturing trends in user interest changes, and continuously enhancing the precision of the recommendation service.

As the main body, libraries should focus on introducing machine learning and artificial intelligence technologies when

building a personalized recommendation service system, training models through user behavior data, and developing precise and efficient recommendation systems, fully utilizing the value of user data in service innovation.

6.3 Collection Resource Decision Support System

The collection resource decision support system based on user behavior data is an important means to achieve precise optimization of collection resources. By analyzing user browsing, borrowing, favoriting, and evaluation behaviors in detail, libraries can accurately grasp the real needs and interest directions of readers, providing scientific decision-making bases for the acquisition and management of collection resources.

In practical work, by analyzing a large amount of user search data and resource utilization data, the types and trends of popular resources can be accurately predicted, and the acquisition plans and collection layout can be optimized in advance. At the same time, by evaluating user feedback on resources, the satisfaction of specific resources can be clarified, and the collection structure and configuration plans can be precisely adjusted to avoid resource waste and significantly improve the utilization efficiency of resources.

In addition, libraries should also establish real-time data analysis and decision support systems, by analyzing the changing trends and hot resource demands of user behavior data in real-time, to adjust collection resource strategies in a timely manner, ensuring the timeliness and effectiveness of resource acquisition and management decisions. At the same time, a unified data governance platform should be built to ensure that data analysis results can be quickly applied to actual collection management [13].

7.Application Strategies and Suggestions

7.1 Construction of Data Governance System

In response to the current deficiencies in data governance within libraries, a comprehensive data governance system should be constructed, mainly including the following aspects: First, establish unified data standards, clarify specific standards for data collection, storage, processing, and application, to avoid data islands and difficulties in utilization due to inconsistent data standards. Second, strengthen data quality management, establish data quality monitoring and evaluation mechanisms, and ensure data accuracy and reliability through continuous data quality checks and optimization. Third, build a data governance platform to promote centralized data storage and unified management, eliminate data islands, and achieve cross-system data sharing and linkage. Fourth, strengthen data security management, establish data access permission management and data security monitoring mechanisms, to ensure the security of user privacy data and sensitive data.

In practice, libraries should first formulate a comprehensive data governance plan, clarify the data governance responsibilities of each business department, and determine detailed processes and standards for data collection, storage, and sharing. In addition, a data governance team should be established, with clear division of responsibilities, and individuals responsible for data quality monitoring, standard implementation, and data maintenance work, to ensure the efficient operation of the data governance system.

7.2 Construction of an Open and Integrated Knowledge Service Platform

To effectively promote data-driven service innovation in libraries, an open and integrated knowledge service platform should be established to achieve effective integration and application of library data with internet data. First, the platform needs to have a high degree of openness and compatibility, capable of seamless connection and data interaction with external internet data platforms; second, by providing API interfaces, it facilitates external developers and researchers to call and analyze data, achieving widespread data resource sharing and deep integration.

Specifically, libraries can build a unified data integration platform, connect and share data with internet platforms such as Douban, Amazon, Dangdang, etc., through data interfaces, and obtain external platform user behavior data in real-time for library service innovation. At the same time, a platform data analysis and visualization tool should be established to provide library staff with convenient data mining and analysis services, enhancing resource management and service decision-making efficiency. In addition, the platform can provide users with personalized resource retrieval and push services, improving the accuracy of user services.

By building an open knowledge service platform, libraries can continuously access external user data resources, form a

service environment with a high degree of integration of internal and external data, and effectively promote the in-depth implementation of data-driven services. In addition, it is important to note that the system platform should strengthen the real-time nature and security management of data, ensuring data utilization efficiency and security, establish unified user privacy protection standards, win user trust, and enhance the sustainability and competitiveness of the service platform.

8.Conclusion and Outlook

This article comprehensively analyzes the application value of internet user behavior data in library service innovation, clarifying its specific role in constructing reader profiles, optimizing collection resource allocation, and precise recommendation services. By analyzing internet user behavior data, libraries can gain an in-depth understanding of user needs and interest trends, thereby providing precise data support for resource acquisition, collection management, and service optimization. Specifically, this article proposes strategies such as building a data governance system and establishing an open knowledge service platform, effectively addressing current issues such as inconsistent data standards, data silos, and insufficient data security, and promoting the efficient integration and deep application of data.

In the future, with the further development of big data, artificial intelligence, and other technologies, libraries will face broader development opportunities in the utilization of internet user behavior data. First, libraries should further explore advanced data fusion technologies, open up channels for data integration within and outside the library, achieve cross-system integrated analysis of different data sources, and establish a more comprehensive and precise data service system. Second, they should actively introduce artificial intelligence and machine learning algorithms, strengthen data analysis capabilities, enhance the precision of reader profiles, the real-time nature of personalized recommendations, and the effectiveness of decision support.

It is worth noting that libraries should strengthen research on user privacy protection and data security, ensuring user privacy and data security during the data application process, and enhancing the compliance and transparency of data governance. By deeply mining user interests and needs trends, libraries can construct a more dynamic and intelligent knowledge service model, truly realizing the intelligent and precise transformation of library services.

In summary, the application of internet user behavior data brings new development opportunities and technological means to library service innovation. Future research needs to pay more attention to the technical implementation of data governance, the optimization of intelligent recommendation algorithms, user privacy protection, and in-depth discussions on data ethics issues, promoting the continuous development of library services towards a more intelligent, precise, and user-needs-oriented direction, comprehensively enhancing the competitiveness and service value of libraries in the digital economy era [14].

Funding

National Library of China's research project "Research on the Application of Internet User Behavior Data in Library Big Data Analysis" (NLC-KY-2018-04)

Conflict of Interests

The author(s)declare(s) that there is no conflict of interest regarding the publication of this paper.

References

[1] Xi J P. Seizing opportunities and proactively implementing the national big data strategy to accelerate the construction of Digital China. CPC News, 2017-12-09. http://cpc.people.com.cn/

[2] Gallagher J, Bauer K, & Dollar D M. Evidence-based librarianship: Utilizing data from all available sources to make judicious print cancellation decisions. Library Collections, Acquisitions and Technical Services, 2005, 29(2): 169–179.

[3] Cheng L J. The practice and implications of big data applications in the U.S.—A library perspective. Information and Documentation Services, 2013, 34(5): 110–112.

[4] Chen H L, Doty P, Yu J C, et al. Library assessment and data analytics in the big data era: Practice and policies. Proceedings of the Association for Information Science and Technology, 2015, 52(1): 1–4.

[5] Fan W H, Li C H, Zhang X W, et al. What kind of "big data" do libraries need? Library Journal, 2012, 31(11): 63–68.

[6] Han C F. The impact and challenges of big data for libraries. Library and Information, 2012, (5): 37–40.

[7] Liu W, Xia C J, & Zhang C J. Big data and linked data: An upcoming data technology revolution. New Technology of Library and Information Service, 2013, (4): 2–9.

[8] Li Y K. Statistical methods for user profiling in the context of big data. Master's thesis. Beijing: Capital University of Economics and Business, 2016.

[9] Liu Y. The application status and prospects of folksonomy. Modern Economic Information, 2010, (5): 205–206.

[10] Chen L B. Folksonomy and its application in libraries. Journal of the Party School of Sichuan Provincial Committee of the CPC, 2010, (2): 93–96.

[11] Wang S, & Xu X. Establishing structured folksonomy based on user classification tags. Library Science Research, 2011, (9): 73–76.

[12] Yang F. Library big data practice based on user profiling: A case study of the National Library of China. Library Forum, 2019, 39(2): 58–64.

[13] Technology Innovation and Social Responsibility Bureau. Notice on accelerating the digital transformation of state-owned enterprises. (2020-07-21) [2023-03-15]. http://www.sasac.gov.cn/

[14] The State Council of China. Notice on the '14th Five-Year Plan' for Digital Economy Development. (2021-12-12) [2023-03-15]. https://www.gov.cn/