# Exploration and Practice of Data-driven Precision Services for Literature and History Disciplines: A Case Study of Beijing Normal University Postprint

**Authors:** Shuning Li

**Date:** 2023-08-27T00:00:00+00:00

## Abstract

[Purpose/Significance] This study explores precision services for literature and history disciplines by leveraging theories and methods of data-driven decision-making.

[Method/Process] By referencing theories of data-driven decision-making and the requirements of precision services for literature and history disciplines, a data-driven service implementation framework is constructed. Data including library borrowing records, access control logs, research outputs, survey responses, user consultation records, and feedback are systematically collected to guide disciplinary service practice through data analysis results.

[Results/Conclusion] Through data analysis, loyal and key users of the discipline are identified, the characteristics and needs of literature and history users' information utilization behavior are understood, critical issues to be addressed during the initial stage of disciplinary services and users' needs for support in teaching, research, and management are identified, and precision disciplinary service strategies are proposed based on these findings.

## Full Text

### Data-Driven Precision Subject Service for Literature and History: A Case Study at Beijing Normal University

**Li Shuning** Beijing Normal University Library, Beijing 100875

**Abstract**

[**Purpose/Significance**] This study explores precision subject services for literature and history disciplines using theories and methods of data-driven decision-making. [**Method/Process**] Drawing on data-driven decision-making theory and the specific needs of precision subject services in literature and history, we constructed an implementation framework for data-driven services. We systematically collected library circulation data, access control data, research output data, questionnaire survey data, user consultation data, feedback data, and other relevant information to guide subject service practice through data analysis. [**Result/Conclusion**] Through data analysis, we identified loyal users and key users, understood the characteristics and needs of literature and history users' information-seeking behavior, identified critical problems to solve during the initial service period, and determined user needs regarding teaching, research, and administrative support. Based on these findings, we propose strategies for precision subject services.

**Keywords:** data-driven decision-making; subject service; literature and history

---

## Introduction

Subject services, as a type of library service, can be traced internationally to the "tracking services" initiated by European and American universities in the late 1970s, and domestically to the establishment of the subject librarian system at Tsinghua University Library in 1998. Over the past two decades, many libraries in China have developed various types of subject services, playing significant roles in disciplinary resource development, teaching and research support, and management decision-making. However, our investigation reveals that successful practices have primarily focused on science and engineering institutions (such as libraries at Tsinghua University, Chinese Academy of Sciences, and Shanghai Jiao Tong University), and even within universities, in-depth subject services have concentrated on science, engineering, agriculture, and medical disciplines. In contrast, subject services for literature and history disciplines have been less effective, even at prestigious liberal arts universities. This disparity arises because humanities researchers exhibit high degrees of "personalized" research and varying levels of information literacy, necessitating effective implementation strategies tailored to the characteristics of literature and history disciplines.

Beijing Normal University Library (hereinafter referred to as "BNU Library") launched its subject service for the School of Chinese Language and Literature and the School of History in late September 2015. Without prior institutional experience or many successful industry cases to reference, the subject service team (hereinafter referred to as "the team") leveraged data-driven decision-making theories to achieve a rapid "ice-breaking" in subject service delivery. By starting with relevant data analysis, the team exploratorily provided precision services to users and achieved positive results.

## 1. Data-Driven Decision-Making Theory

The concept of data-driven decision-making generally originated in the field of education in the late 1970s, primarily using student academic performance and classroom behavior data to improve decision-making quality in basic education [5-6]. Although the concept has existed for nearly 40 years, its application was limited due to technological constraints and practical needs. While used in business, its scope remained narrow until the concepts, technologies, and methods of big data gained prominence after 2012, leading to broader adoption across more fields and domains.

Currently, there is no unified definition of data-driven decision-making. Some define it as "the practice of making decisions based on data analysis rather than relying solely on intuition" [7], while others describe it as "a form of decision-making where decision-makers base their choices on information and evidence obtained through relevant data analysis" [6]. Yale University scholar S. Tudesco defines it for libraries as "the thinking activity and work process where librarians form decision recommendations and solutions based on objective data mining and analysis rather than personal experience and subjective speculation" [8]. These definitions reveal that data-driven decision-making emphasizes basing decisions on objective data, using information and knowledge obtained through data analysis to replace decision-makers' bounded rationality, thereby enhancing the scientific nature of the decision-making process.

The operational principle of data-driven decision-making is based on the pyramid model of "data—information—knowledge" [9]. People obtain relevant data through observation, measurement, and research of objects, then organize and integrate the collected data to produce useful information, and further process the information to obtain instructive knowledge. Under this principle, the operational mechanism of data-driven decision-making theory includes: goal setting, data collection, information mining, information presentation, and decision evaluation. Data-driven decision-making is mainly implemented through four approaches [6]: (1) judging the future through "history" to improve long-term and sustained decision-making performance; (2) making decisions based on "evidence" to enhance scientific and accurate decision-making performance; (3) verifying decisions through "simulation" to improve practical decision-making performance; and (4) evaluating decisions through "tracking" to make decisions more comprehensive and refined.

In recent years, data-driven decision-making theory has been increasingly applied to library fields both domestically and internationally, covering various aspects of resources, services, technology, and management. For example, the Third Military Medical University Library used circulation data to determine print book copy numbers for different disciplines [10]; the National University of Singapore Library and Singapore Management University Library set flexible

loan periods of 3 hours to 28 days for books based on user reservation data; East Tennessee State University in the United States used socioeconomic data from students' pre-enrollment regions to determine content for library information literacy courses [11]; Brigham Young University Library in the United States used web2py to achieve completely user data-driven website design and development [12]; and Oxford University's Bodleian Library used user data to make staffing decisions for its dense book stacks [13]. Data-driven decision-making has increasingly become a research and practice hotspot in the library and information science field both at home and abroad.

---

## 2. Needs and Implementation of Precision Subject Service for Literature and History

To achieve an "ice-breaking" in subject service, it is essential first to grasp the precision needs at the overall user level, then progressively explore the precision needs of individual user groups. Therefore, during the initial stage, the primary challenges facing the team were: how to quickly identify loyal and core users, rapidly understand the characteristics of literature and history users' literature utilization, and promptly comprehend the universal precision needs of these users. Given the importance of user behavior data in the big data environment for subject services, the team decided after repeated discussion to leverage library data-driven decision-making theories and start with relevant data analysis to construct an implementation framework for data-driven subject service for literature and history (see Figure 1 [Figure 1: see original paper]).

This implementation framework uses relevant data analysis tools (such as Excel and CiteSpace) to analyze various types of user utilization data and directly guides subsequent practice with the analysis conclusions. The right side of the framework lists some key data required for data-driven decision-making in literature and history subject services, such as user circulation data, digital resource utilization data, user consultation data, user survey and interview data, and citation data from user research outputs. These represent only partial relevant data that can be supplemented with other data as needed based on practice and data collection conditions. The left side of the framework outlines the work content of subject services. During the initial service period, three key points must be addressed: (1) identifying loyal and core users and establishing connections with them; (2) addressing one or several key problems users encounter in library literature utilization; and (3) understanding the characteristics of users' literature utilization behavior.

Comprehensive subject services provided by university libraries both domestically and internationally generally fall into five categories: (1) subject liaison, building a bridge between the library and departmental users; (2) subject resource development, assisting in disciplinary resource construction based on user needs obtained through questionnaires and interviews, as well as collection de-

velopment policies; (3) teaching support, providing teaching materials support or assisting in information literacy education according to disciplinary users' teaching needs; (4) research support, providing support for disciplinary users' scientific research, such as assisting in locating or providing needed literature and recommending journals for submission; and (5) management support, providing data support for disciplinary management and evaluation work, such as disciplinary development trends and comparative analysis of research output with peer institutions. As subject services develop in depth, generalized "surface-level" services can no longer meet user needs and expectations. More libraries and users hope to improve service targeting and precision to achieve multiplier effects. The content of literature and history subject services is no exception. Moreover, due to the high degree of "personalized" research among literature and history scholars and the scarcity of relevant research and cases, grasping precision needs is particularly challenging. It should be noted that precision needs exist at two levels: one is the precision needs of literature and history users as a whole, distinct from other disciplines; the other is the precision needs of individual users (or groups) within literature and history, distinct from other users (or groups) in the same discipline. As the second-phase subject service launched in BNU Library's Subject Service 3.0 version, it was crucial to quickly complete the "ice-breaking" journey and provide precision services to users based on objective data analysis.

---

## 3. Data-Driven Decision-Making Practice in BNU Library's Literature and History Subject Service

### 3.1 Identifying Loyal, Core, and Potential Users

**3.1.1 Data Analysis Methods** To understand users, it is first necessary to establish a user dataset for service objects, which proves highly effective. According to BNU Library's focus on graduate students and above in this round of subject services, the team established datasets for faculty and graduate student users. The specific process involved collecting and organizing faculty names, disciplines, courses taught, research directions, research projects, and contact information from the two schools' websites, supplemented by faculty rosters obtained from the schools to create the faculty user dataset. Graduate student basic information was obtained from the library's integrated management system to form the student user dataset.

**(1) Mining loyal users.** Using the library's integrated management system, the team analyzed book borrowing data from faculty and graduate students in the two schools over the past five years. They identified users with continuous borrowing behavior each year whose borrowed books were basically for research purposes (excluding users who borrowed books for family children's education, as discovered through data analysis). Additionally, referencing access records for faculty and students from the two schools in the library's access control system,

users who frequently visited the library and borrowed books were identified as loyal users.

**(2) Mining core users.** The subject service team asked the library director to contact relevant departmental leaders to convene user survey symposiums. Participants included deans, associate deans, directors of major departments, and academic backbones. These users more easily grasp the macro-level disciplinary development direction and overall needs, representing important supporting forces for successful subject service implementation. These individuals, along with other academic leaders, academic backbones, and some influential loyal users, constitute core users. The disciplinary user dataset serves as an important pathway for mining other core users.

**(3) Mining potential users.** These users should have strong needs in their teaching and research but rarely use the library, such as those who have recently obtained new research projects, opened new courses, or joined as new faculty. Logically, these users have urgent needs for literature resources. The subject service team first used the aforementioned disciplinary user dataset to mine some potential users, with further identification of other potential users as subject services deepen.

**3.1.2 Service Strategies** For different user groups, the subject service team adopted a tiered approach [18], promoting services, establishing connections, and soliciting needs and opinions through different channels and starting points, then providing follow-up services. For loyal users, the team established connections by regularly emailing new book lists relevant to them, accompanied by promotional materials about team services. For core users, the team conducted on-site interviews, starting with questions like "whether the library's collection of your published monographs needs supplementation" to introduce team work content and plans. For potential users, the team combined email inquiries with some on-site interviews, using congratulations on obtaining important projects as a reason to promote team services and establish connections.

### 3.2 Analyzing User Literature Utilization Needs and Behavior Characteristics

**3.2.1 Data Analysis Methods** The team primarily analyzed user literature utilization needs through questionnaire surveys, symposiums, interviews, consultations, and opinion feedback data from various channels (including the director's mailbox, university BBS, WeChat, and Weibo). Citation analysis of users' published works was used to analyze their literature utilization behavior characteristics, with the two data sources complementing each other. To avoid "bias" in questionnaire design by librarians, the team invited doctoral students from the disciplines to independently design, distribute, collect, and analyze questionnaires according to the team's research needs.

**(1) Analyzing user literature utilization behavior based on citation data.** The team analyzed over 60,000 citation data entries from literature and history faculty and students' research outputs published in CSSCI journals between 2004 and 2014, revealing characteristics of literature utilization behavior: (1) The average number of citations per article in literature and history is more than twice that of science and engineering disciplines, indicating heavy reliance on literature resources. Simultaneously, literature and history research features distinct "individual research and creation" characteristics, with a considerable portion of research outputs having zero citations (14% for history, 17% for Chinese literature) and a large proportion of self-citations. (2) Literature and history users employ diverse document types, including books, journals, newspapers, archives, manuscripts, oral materials, and images (over 20 types), but more than two-thirds of needs concentrate on Chinese and foreign books, with weaker demand for journals (journal citation ratios are only 14% for history and 27% for Chinese literature). (3) Chinese literature is the primary language type used by literature and history users (91% for Chinese literature discipline, 93% for history), with overseas Sinology literature being the focus of foreign language materials. (4) The citation half-life for literature and history literature exceeds 20 years, showing particular attention to classic older documents. (5) Literature and history researchers pay special attention to specific editions from specific publishers when citing books. (6) Although literature and history researchers share similar literature utilization behaviors, significant differences exist and require differentiated treatment.

**(2) Analyzing needs characteristics through demand data.** Through organizing demand data collected from various channels, the team identified significant characteristics of literature and history user needs [19]: (1) Literature and history users' teaching and research are highly dependent on literature resources, with resource needs far exceeding needs for other subject service content. Resource needs extend beyond the library's collection, with strong demand for books from the National Library of China and free literature and history resources on the internet. (2) Literature and history users require diverse resource types but have far greater needs for books than other types, Chinese literature than foreign literature, and older resources than new resources. (3) Although literature and history users strongly prefer print resources, they also have significant needs for digital resources. Various departmental research groups have collected and digitized many documents, but sharing is limited to the same mentor or research group, requiring expanded common knowledge and sharing scope. (4) Open access resources are scarce in literature and history, and retrospective literature is difficult to obtain, especially documents from 1949-1980 in China, which represent a "blind spot" in digital resource construction for literature and history, with print resources also difficult to provide. (5) For resource acquisition, literature and history users prefer direct access and pay less attention to acquisition channels, methods, and processes. (6) Literature and history users' needs for subject services are highly personalized, making universal service content and strategies difficult to apply. Subject services need

to incorporate more "humanistic care."

Additionally, the user needs survey revealed four noteworthy findings: (1) The library website remains an important portal for literature and history users to access literature resources, with 44% of surveyed users visiting daily and another 21% visiting 3-5 times weekly. (2) Among general databases, Duxiu Chinese Academic Search is the second most used Chinese database by literature and history users after CNKI, with its "document delivery" function being highly favored, far exceeding the library's document delivery service. The China Basic Ancient Books Database is the most used specialized database. (3) The most used foreign language database by literature and history users is the CALIS Foreign Language Journal Network, whose Chinese interface is an important reason for user selection, as interface language affects specific usage. (4) For catalog and chapter title databases of ancient books, local gazetteers, and other literature resources without full text, literature and history researchers maintain strong interest.

**3.2.2 Service Strategies** Based on needs and behavior characteristics obtained through data analysis, the team mainly carried out the following work: (1) Addressing users' heavy reliance on book resources, the team sorted out new library collections closely related to users' majors and research directions based on the user dataset and promoted them via email. Focusing on urgently needed overseas Sinology books, the team sorted existing collections, compiled them into thematic catalogs sent to relevant faculty and students, and held online exhibitions of overseas Sinology collections to promote available resources. Simultaneously, the team compiled catalogs of recent or classic overseas Sinology resources not collected by the library, asked subject teachers to select items, and conducted centralized supplementation. (2) Using the subject resource promotion month opportunity, the team provided training on frequently used Chinese and foreign databases selected by users. Even for user-friendly databases like CNKI and Duxiu Search, the team conducted in-depth training focusing on important but overlooked functions. The team also analyzed reasons for low usage of databases that users should be using but weren't, conducted targeted promotion, and used the subject service platform to centrally display high-quality open access databases (including title databases) and important databases available to literature and history users from other institutions like the National Library of China. (3) The team fed back to the library's circulation department to prioritize support for book literature from 1985-2000 heavily used by literature and history users according to citation half-life patterns. The team also requested the acquisition department to contact vendors for trial access to databases covering 1949-1980 literature, such as the "Guangming Daily" full-text database developed by the Green Apple Company. (4) The team served as an intermediary for electronic resource sharing between research groups, collecting and compiling simple electronic catalogs of copyright-free resources mastered by groups willing to share, then sharing them with other needed faculty and students. (5) Understanding the distribution of literature needed for key users' research

directions, the team proactively helped faculty and students obtain difficult documents through interlibrary loan and document delivery services, using this opportunity to promote these services. (6) The team increased humanistic care in subject services, visiting senior professors and scholars from literature and history schools such as Liu Jiahe, Wang Ning, and Nie Shiqiao before and after holidays to promote subject services, provide personalized resources, and solicit valuable donations. (7) The team cooperated with literature and history school teachers to offer practical courses in "Philology." Library ancient book experts explained the physical form and binding formats of ancient books using direct presentation methods and led students to visit library collections of imprints and manuscripts from various periods to experience different binding formats and multi-color printing, woodblock, and various movable type techniques.

### 3.3 Addressing Key Service Problems and Other Subject Service Work

**3.3.1 Data Analysis Methods** In addition to the aforementioned service work centered on literature and history users' utilization characteristics and needs, the team also used citation data from faculty and graduate students' published works and theses to analyze collection guarantee rates and conduct necessary supplementation, which will not be elaborated here. The team collected problems urgently needing resolution during the initial subject service period through various channels of opinion feedback data, using book circulation, book recommendation, document delivery, user consultation, and research output data in relevant subject areas to support resource development and teaching and research.

**3.3.2 Service Strategies** During the initial work period, to increase user trust and dependence, the team needed to focus on solving one or two problems that subject users considered difficult. Through analysis of previous user opinion feedback data, the team discovered that "literature and history users' request to increase concurrent users for the China Basic Ancient Books Database" was an urgently needed problem. This issue was mentioned multiple times in problem feedback, and a joint letter from dozens of professors had been received in the previous year. The library had actively applied for funding to increase concurrent users. After the team's intervention, they first investigated concurrent user numbers for this database at peer institutions and carefully analyzed the database's usage data, discovering that the real problem was not "too few concurrent users" but rather that "the system setting for 'forced release of users who occupy the system without using it for a long time' was too long." The team successfully solved the problem by adjusting the relevant time settings, earning recognition from many users and saving the library hundreds of thousands in literature funds originally planned for purchasing concurrent users.

Additionally, the team carried out the following work based on data: (1) Addressing the difficulty of selecting foreign language books for the discipline,

the team used the disciplinary user information dataset and book circulation records to identify users who frequently used foreign language books and loyal users who regularly recommended books and used document delivery. Considering disciplinary balance, these users were designated as key promotion targets for foreign language book selection, achieving good results. (2) The team organized and analyzed user consultation data to develop information literacy course content with different focuses for undergraduate and graduate students based on their primary consultation questions, shifting from teaching based on librarians' understanding to teaching based on user concerns, thereby improving relevance and effectiveness. (3) Addressing graduate students' difficulty in selecting journals for publishing Chinese papers, the team compiled a list of journals where the school had published papers using CNKI data, ranking them by publication volume. The top 20 journals, along with contact information and submission guidelines, were provided to graduate users. For faculty difficulty in selecting journals for foreign language articles, the team compiled SSCI and A&HCI journals where mainland Chinese scholars in relevant disciplines had published most frequently, ranking them by publication volume and selecting the top 30 journals. Journal profiles, impact factors, contact information, and submission guidelines were compiled and provided to relevant faculty users, earning widespread user recognition.

## Conclusion

This paper designed a specific implementation framework for literature and history subject services based on data-driven decision-making theory and introduced BNU Library's concrete practice, exploring some referential experiences: (1) Literature and history users' literature utilization habits are indeed very different from those of science, engineering, agriculture, and medical users. For example, their need for literature resources themselves exceeds their need for information literacy education on how to obtain resources; they have strong demand for document literacy content and practice such as ancient book binding and edition identification; they reference diverse literature types but have far greater needs for books than other types, Chinese literature than foreign literature, and older resources than new resources; they show particular preference for books from key publishers and classic Chinese works, while focusing on overseas Sinology materials for foreign resources; the discipline's literature half-life is long; and subject services need to incorporate more humanistic care. These aspects deserve special attention from literature and history subject service providers. (2) It is crucial for team members to establish a data-driven decision-making concept, grounding service starting points in objective data analysis rather than subjective experience and speculation, which greatly benefits subject service effectiveness. Using user behavior data and survey data to comprehensively understand disciplinary users' behavior characteristics and needs is a critical step in subject service development. Although data processing and analysis consume substantial time and effort, understanding users' literature utilization characteristics through citation analysis of institutional users'

academic outputs is more systematic and objective than questionnaire surveys and interviews. Having users design and conduct surveys from their own perspectives can yield unexpected insights. Utilizing user data enables subject service work to follow clear procedures, such as identifying loyal and core users, and using consultation record data to identify and solve key problems, which is very helpful for establishing trust and dependence with users.

Of course, this paper only discusses the data-driven decision-making perspective for developing precision literature and history subject services, and the selected cases are only typical examples and partial attempts during the team's ice-breaking journey. There remains room for further exploration in data-driven decision-making applications, including cost-effectiveness analysis of literature resource utilization and evaluation of resource services based on data. Although other universities may not explicitly conduct research and practice under the name of "data-driven decision-making," they also have some practices using various data analyses to carry out work [16-17], and research results have begun to consider university library subject services from a data perspective [18], which are also worth further attention and reference.

---

**References**

[1] Liu Ying. Reflections on the positioning and services of humanities and social sciences subject librarians[J]. Library and Information Service, 2012, 56(S2): 48-50.

[2] Yuan Yi. Discussion on subject services in social science research professional libraries[J]. Information and Documentation Services, 2007(6): 88-91.

[3] Gerile, Na Min. Reflections on university library service concepts based on humanities users' needs[J]. Journal of Inner Mongolia Agricultural University (Social Science Edition), 2009, 11(4): 215-216.

[4] Feng Fang. Research on embedded subject services in university libraries for ancient literature and history specialties[J]. Journal of Lanzhou Institute of Education, 2014, 30(12): 68-69.

[5] Morrissey L. Data-driven decision making in electronic collection development[J]. Journal of Library Administration, 2010, 50(3): 283-290.

[6] Chang Tongshan. How to improve university decision-making performance: Institutional research and data-driven decision-making model perspective[J]. Fudan Education Forum, 2013, 11(2): 54-60.

[7] Irving W. Data-driven decision making: promises and limits[EB/OL]. [2018-04-03]. http://blogs.wsj.com/cio/2013/09/27/data-driven-decision-making-promises-and-limits/.

[8] Tudesco S. What is a data-driven academic library[EB/OL]. [2017-10-26]. http://www.slideshare.net/studesco/data-driven-libraries20131204.

[9] Huang Chen. Research on public security decision-making methods based on "data-driven decision-making" theory[J]. Beijing Police College Journal, 2014(6): 49-55.

[10] Zhang Jingli, He Xiaoyang. Enlightenment of book procurement based on circulation data analysis: a case study of the Third Military Medical University Library[J]. Lantai World, 2017(22): 78-80.

[11] Leslie A, Mark E, Kathy C, et al. Redirecting library instruction based on socioeconomic data[J]. Reference Services Review, 2010, 38(3): 398-416.

[12] Mathew M. Using web2py Python framework for creating data-driven web applications in the academic library[J]. Library Hi Tech, 2016, 34(1): 164-171.

[13] Mathew W. Using data to strategically deploy staffing resources[J]. Library Management, 2018, 39(3/4): 166-171.

[14] Li Shuning. Constructing a tiered subject service system to provide precision embedded services[J]. Library Work and Research, 2017, 1(1): 85-89.

[15] Yu Jing, Liu Yingchun, Li Shuning, et al. Exploration and practice of subject services based on disciplinary characteristics and user information behavior analysis: a case study of BNU Library's Subject Service 3.0[J]. Journal of Academic Libraries, 2017, 35(5): 80-85.

[16] Zeng Xiaomu, Lin Jia. Practice of subject services based on effective disciplinary information behavior: a case study of Tsinghua University's mathematics discipline[J]. Journal of Academic Libraries, 2014, 32(3): 85-89.

[17] Li Shen, Li Ji. Precise positioning of subject service goals and precision marketing of subject services[J]. Library Science Research, 2013(9): 79-81.

[18] Wang Hongbo. Research on precision subject services in university libraries driven by small data thinking[J]. Hebei Sci-Tech Library Journal, 2018, 31(1): 44-47.

*Note: Figure translations are in progress. See original paper for figures.*