Usage Statistics and Analytical Reflections on Foreign Language Database Resource Construction in Academic Libraries: A Case Study of Four Databases at Fudan University (Postprint)

Authors: Zhao Weijia, Hao Qun, Libin Zhang

Date: 2023-04-01T15:51:23+00:00

Abstract

[目的/意义] Since the 21st century, the proportion of databases in university library procurement resources has been continuously increasing, with foreign-language databases representing a crucial source of electronic resources. Concurrently, in the face of stringent requirements from various supervision and audit processes, library staff have encountered numerous challenges during the procurement and development of foreign-language database resources. Taking Fudan University Library as a case study, this research analyzes and investigates the resource subscription and utilization patterns of four subscribed foreign-language databases—Elsevier SD, Wiley, Springer, and Taylor-ST—to provide insights and references for decision-making in foreign-language database development in university libraries.

[方法/过程] Based on statistical data from the four foreign-language databases, this study categorically summarizes the overview of collection resources, the support status for "Double First-Class" discipline construction, annual utilization patterns, and the annual status of highly-cited journals not held in the collection. From these four analytical perspectives, data is coded using grounded theory methodology, combined with specific data to generate a comprehensive usage table for foreign-language databases, and the grey relational analysis method is employed to calculate and evaluate the influence degree of each indicator.

[结果/结论] Through computational analysis, five key indicators most significantly influencing the usage of foreign-language databases at Fudan University Library are identified: annual download volume, citation volume, publication volume, total journal holdings, and proportion of core journals. From three perspectives—supplementary metrics for academic achievements, bibliometrics, and resource provision—optimization recommendations are proposed for the development of

foreign-language databases in Chinese university libraries, offering methodological references and establishing targets, basis, and strategic directions for key control points in subsequent research on information resource development at other institutions.

Full Text

Preamble

The Usage Statistics and Analytical Thinking of Foreign Language Database Resource Construction in University Libraries: Taking the Practice of Four Databases at Fudan University as an Example

Zhao Weijia[1], Hao Qun[1], Zhang Libin[2]
[1]Fudan University Library, Shanghai 200433
[2]Nankai University Library, Tianjin 300071

Abstract: [Purpose/Significance] Since the 21st century, the proportion of databases in university library procurement resources has been continuously increasing, with foreign language databases serving as an important source of electronic resources. Meanwhile, facing the strict requirements of various supervision and auditing processes, staff have encountered numerous problems in the procurement and construction of foreign language database resources. Taking Fudan University Library as an example, this study analyzes the resource ordering and usage conditions of four foreign language databases (Elsevier SD, Wiley, Springer, and Taylor-ST) to provide ideas and references for decision-making in foreign language database construction in university libraries. [Method/Process] Based on the statistical data from the four foreign language databases, this paper categorizes and summarizes the overview of collection resources, the guarantee overview for "Double First-Class" discipline construction, annual utilization conditions, and annual highly-cited journals without holdings. From these four analytical perspectives, data was coded based on grounded theory, and a comprehensive usage table of foreign language databases was obtained through specific data. The grey relational analysis method was used to calculate and evaluate the influence degree of each indicator. [Result/Conclusion] Through calculation, five key indicators were identified that most significantly affect the usage of foreign language databases at Fudan University Library: annual download volume, citation volume, publication volume, total number of journals in the collection, and proportion of core journals. From three perspectives—supplementary measurement of academic achievements, bibliometrics, and resource guarantee—optimization suggestions are proposed for the construction of foreign language databases in Chinese university libraries, providing methodological references and key targets, basis, and ideas for subsequent research on information resource construction in other universities.

Keywords: university library; foreign language database; resource construction; grounded theory; grey relational analysis

Classification Number: G250
DOI: 10.13266/j.issn.0252-3116.2022.04.008

With the construction of digital and smart libraries, users' demand for and frequency of using electronic resources have greatly increased, and the focus of library literature resource construction has gradually shifted from print to electronic resources. As the main body representing universities in database procurement, university libraries' expenditure on electronic resources occupies a "major portion" of annual funding. According to the "2019 Statistics on Electronic Resource Procurement Expenses of 1,165 University Libraries" [1], "2019 Statistics on Literature Resource Procurement Expenses of 1,263 University Libraries" [2], and "2019 Statistics on Annual Total Funding of 1,278 University Libraries" [3] released by the Higher Education Library and Information Work Guidance Committee of the Ministry of Education, the top three university libraries in terms of electronic resource procurement expenses—Tsinghua University Library, Peking University Library, and Sun Yat-sen University Library—spent 52,038,548 yuan, 51,809,944 yuan, and 50,536,164 yuan respectively, accounting for 74.79%, 69.63%, and 74.59% of their literature resource procurement expenses. The Chinese University of Hong Kong Library also mentioned in its 2019-2020 annual report [4] that electronic resource expenses accounted for 92% of material expenditures. Table 1 shows the numbers of foreign language and Chinese electronic databases purchased by eight comprehensive universities ranked in the top 50 of the 2022 QS World University Rankings.

As seen in Table 1 , among these eight institutions, the number of foreign language databases purchased accounts for more than half of the total. Considering the important influence and exemplary role of these eight comprehensive universities in Chinese academic research, analyzing the usage of foreign language databases and evaluating the effectiveness of purchased foreign language databases is necessary to provide suggestions for future foreign language database resource construction in libraries. In view of this, this study combines existing research with the actual usage of foreign language databases in Chinese university libraries (taking Fudan University Library as an example). By collecting and analyzing resource attribute data and utilization data of current foreign language databases, this paper codes the data based on grounded theory and uses grey relational analysis to identify key indicators, thereby proposing optimization strategies and suggestions for foreign language database resource construction in university libraries from the dimensions of resource guarantee, supplementary measurement of academic achievements, and bibliometrics.

2 Literature Review

2.1 Theoretical Review of Problems and Solutions in Foreign Language Database Resource Construction

As an information repository for learning about the latest foreign technologies and scientific developments, libraries currently purchase foreign language

databases mainly through two methods: whole-database purchase or purchase by album/subject. Existing foreign language database construction mainly faces problems such as funding pressure, resource structure issues, lack of initiative in resource construction, resource ownership issues [13], and how to evaluate the construction status of foreign language database resources. Regarding funding pressure, Wang Chunsheng [14] believes it is mainly caused by domestic university libraries' incomplete understanding of data vendors' sales and pricing strategies. Therefore, libraries should enhance the negotiation capabilities of database procurement personnel to obtain favorable procurement conditions. Regarding unbalanced database resource structure and the lack of initiative to adjust collection resources during contract periods due to consortium purchasing, Zhao Yanzhi [15] suggests that Chinese university libraries can form negotiation project teams, consider changing resource procurement models (such as "pay-per-article + interlibrary loan + open access") when renegotiating bulk transaction terms, and learn from successful practices of foreign libraries in resisting big deals. Regarding database resource ownership issues, libraries should adopt a researcher perspective, investigate and analyze licensing agreement models used in digital resource procurement by university libraries to avoid potential intellectual property disputes and improve the cost-effectiveness of digital resource procurement [17]. On the other hand, university libraries can also break the existing pattern and create a new academic communication model of "scholars create academic achievements → store in institutional repositories or online disciplinary platforms → scholars freely utilize and create new academic achievements" [17] through open access, thereby weakening data vendors' control over the entire academic communication process. In terms of evaluating foreign language database construction, current research methods mainly include database usage performance evaluation [18], evaluation of digital resources from user data dimensions [19], analysis of database usage statistical reports [20], and research on the construction of academic service capability [21] and effectiveness evaluation systems [22] in university libraries.

2.2 Practical Review of Foreign Language Database Resource Construction

In addition to theoretical research by scholars at home and abroad, universities have also applied theory to practice. To improve the foreign resource retrieval experience, Shanghai Jiao Tong University has continuously strengthened the application of foreign full-text electronic journal models, improved resource maps, increased integration of journals and publication/citation resources, introduced big data analysis methods to deeply explore academic literature usage patterns, evaluated the influence of academic books and publishers, and studied citation work of foreign academic literature to promote the transformation of resource construction [24]. Regarding how libraries can promote innovative exploration of foreign language database collection resources, Yang Xin et al. [24] suggest that university libraries should refer to the correlation and clustering characteristics of disciplines and departments in electronic resource utilization

to optimize electronic resource allocation and procurement from multiple dimensions. All 10 campus libraries of the University of California launched a unified UCLibrarySearch retrieval and borrowing system on July 27, 2021 [25]. This retrieval system makes resources (including native and foreign language resources) within the 10 libraries easier to find and borrow through improved interlibrary sharing tools and a simplified user interface, thereby maximizing the utilization of the University of California's collection. Delft University of Technology and the National Library of the Netherlands launched the Future Libraries Lab to investigate user behavior patterns and activities, helping them discover the library's multilingual digital collection resources in newer and more interesting ways [26]. In terms of promoting open access to data resources, many universities have taken action. For example, Oxford University Library's 2022-2024 plan mentions "increasing access to collections and promoting greater discovery and engagement," with key measures including extending library opening hours, expanding the scope and time span of borrowable multilingual electronic collections, and strengthening community participation and cooperation between libraries and museums, archives, and art galleries [27].

In summary, some scholars have analyzed the problems and strategies of foreign electronic journal database procurement in universities, while others have conducted analyses and evaluations of foreign language database resource construction from specific perspectives or theories, and some universities have applied these theories in practice. These provide a good foundation for this study. However, existing research also has some shortcomings: First, although studies have shown correlations between resource usage and research output at the journal [28-29] and institutional [30-32] levels, for Chinese universities, the difference between foreign language electronic journals and Chinese scholars' native language necessitates consideration of their particularities in analysis. For example, when facing non-native literature, users are more accustomed to downloading it locally first and then using translation software to read it; or some foreign electronic journal databases have recommendation functions during download, automatically recommending "similar" articles to users. Due to unfamiliarity with the non-native language and the need for literature breadth, users often accept the system's automatic recommendations and download these similar articles, but whether these articles actually play a role in users' research processes cannot be determined. These processes are not completely transparent but rather more like a "grey system"—the research process of users using foreign language databases is not entirely clear, and the relationship between factors such as journal browsing and download volumes and the actual useful literature categories and quantities is not fully understood. Therefore, browsing and download volumes alone are insufficient to fully indicate usage effectiveness. Second, current strategies used by university libraries in selecting foreign language electronic journal databases often rely solely on single-article download cost for quantitative evaluation, but as mentioned above, this perspective is too singular to comprehensively measure user usage of foreign language databases. To fill these research gaps, this paper first analyzes the resource attribute data of four

foreign language databases to gain a macro understanding of the library's foreign electronic journal database collection, then uses grey relational analysis based on utilization data from these four databases to identify key indicators that should be considered in resource construction from both traditional statistical and statistical perspectives, and finally provides reflections on foreign language database resource construction in university libraries based on the calculation results.

3 Statistical Analysis and Evaluation Practice of Foreign Electronic Journal Databases at Fudan University Library

3.1 Practice Data

All data used in this practice comes from four foreign electronic journal databases ordered by Fudan University Library (Wiley, Taylor-ST, Elsevier SD, and Springer). It should be noted that although the number of selected foreign databases is relatively small compared to the total number of foreign databases available on the market, these four databases collect journals, books, reference works, etc. in various disciplines including environmental and agricultural sciences, chemistry, engineering, computing and engineering technology, physics, mathematics, life sciences, medicine, humanities and social sciences, with large quantities and varieties [33]. Moreover, these four foreign database vendors are internationally recognized high-quality academic publishing companies, and most of their publications are core materials in related disciplines [28-29] with long histories and rich publishing experience. Therefore, selecting these four foreign databases as samples for analysis can to some extent reflect the utilization behavior characteristics of foreign databases in universities. Statistical data dimensions include the usage volume and denial access volume of the four databases in 2019 and 2020, journal data from the four foreign electronic journal databases purchased in 2020-2021, publication data of Fudan University scholars in Web of Science (WoS) from 2018 to present, and citation data in WoS from 2010 to present.

(1) Usage data of the four foreign language databases. By journal unit, fields include full journal title, database vendor, publishing platform, DOI, ISSN, total user downloads, number of downloading user IDs, annual total downloads, and monthly downloads by year.

(2) Denial access data of the four foreign language databases. By journal unit, fields include full journal title, database vendor, subject classification, annual total denial access volume, and monthly denial access volume by year.

(3) Journal collection data of the four foreign language databases. By journal unit, fields include journal number, ISSN, full journal title, database vendor, print price (USD), subject classification, whether online version is available, electronic journal price (USD), 2020 electronic journal price in

RMB, and citation score.

(4) RMB prices (including bundled journals) of the four foreign language databases. By database unit, fields include database name, 2019 pricing, and 2020 pricing.

(5) Publication and citation data of Fudan University scholars. By journal article unit, the original data consists of journal articles published by Fudan University scholars from 2018-2020, with fields including article accession number, cited references (after data processing, references are split article by article and can be matched back to the original article using the accession number field; if the original article has 30 references, it is divided into 30 rows in the reference splitting. Due to Excel row limitations, this list is placed in two sub-tables: reference split 1 and reference split 2, which together constitute all references, totaling more than 130,000 articles. Journal names in references are mostly abbreviations, and a mapping table between full and abbreviated journal names is available for reference conversion), number of cited references, authors, language, document type, abstract, WoS Core Collection citation count, total citation count (WoS Core, Biosis Citation Index (BCI), and Chinese Science Citation Database (CSCD)), usage count (last 180 days), usage count (2013 to present), publisher, ISSN, EISSN, publication date, volume, issue, DOI, page count, WoS category, research direction, open access indicator, etc.

3.2 Analysis of Collection Resources of Four Foreign Language Databases

3.2.1 Subject Distribution of Collection Resources Using Tableau software, we analyzed the collection quantity comparison among the four foreign language databases, the comparison of collection subjects, and the proportion of each subject in each database, presenting the results visually in Figure 1 [Figure 1: see original paper]. The left side of Figure 1 [Figure 1: see original paper] shows the differences in collection volumes among the four foreign language databases, while the right side shows the collection status of different subjects from the four foreign language databases at Fudan University Library. Overall, the most abundant resources are in science, medicine, and engineering. The flow diagram in the middle of Figure 1 [Figure 1: see original paper] shows that in the Elsevier SD database, the subjects with larger collection quantities are science, engineering, and medicine, while there are no holdings in philosophy, education, literature, history, or military science. Springer has the largest collection volume among the four databases, with science being the most prominent subject compared to others in the Springer database, followed by medicine, engineering, and law, while humanities-related subjects such as philosophy, history, and literature have no relevant holdings. Taylor-ST has the smallest total collection volume among the four databases, with holdings mainly distributed in science, engineering, and medicine. In the Wiley database, Fudan University Library has holdings in all subject areas except military science and management, with

the largest quantities in science, medicine, and education.

3.2.2 Quality of Collection Resources By analyzing the journal data included in the four foreign language electronic journal databases purchased by the library in 2020-2021, we can count the total number of foreign language electronic journals in each database, the number of duplicate journals among databases, the net journal proportion, and the proportion of core journals by comparing the collection lists of the four foreign language electronic journal databases with the Journal Citation Reports (JCR) core journal directory [34]. The specific results are shown in Table 2 .

As shown in Table 2 , although the four foreign electronic journal databases differ in the number of journal titles purchased, there are no duplicate purchases among the databases, indicating a relatively reasonable database selection strategy.

3.2.3 Guarantee of "Double First-Class" Discipline Construction According to the "Double First-Class" Construction Discipline List issued by the Ministry of Education of the People's Republic of China, Fudan University has 17 disciplines selected: philosophy, political science, Chinese language and literature, Chinese history, mathematics, physics, chemistry, biology, ecology, materials science and engineering, environmental science and engineering, basic medicine, clinical medicine, integrated traditional Chinese and Western medicine, pharmacy, mechanical and aerospace and manufacturing engineering, and modern linguistics. The corresponding subject categories of these 17 "Double First-Class" disciplines are shown in Table 3 .

By calculating the number of matching journals, journal matching rate, denial access rate, number of core matching journals, and core guarantee rate of the four databases' collection resources corresponding to each subject category, the library can understand the guarantee overview of the four foreign language databases for "Double First-Class" discipline construction.

As shown in Table 3 , from a vertical perspective of foreign language databases, the four foreign electronic journal databases at Fudan University Library each have different emphases: Wiley provides important literature resources in philosophy, literature, and history, while Springer is better in law and engineering. Elsevier SD has the highest core journal guarantee rates in science, engineering, and medicine. In contrast, Taylor-ST is relatively inferior in supporting literature resources in various subjects except science. From a horizontal perspective of each subject, the four databases have larger total collection quantities in science, engineering, and medicine, with science having the highest core journal guarantee rate, followed by medicine and engineering. Philosophy, law, literature, and history have fewer journals in these four databases and correspondingly lower core journal guarantee rates, indicating that the current focus of the library's resource construction in these four foreign language databases is on science, engineering, and medicine.

3.3 User Utilization Data Analysis

3.3.1 Annual Utilization of Four Foreign Language Databases Considering that readers' usage behavior of foreign language databases is based on their own needs and motivations, the usage data generated when readers use foreign language databases can be regarded as a concentrated reflection of their usage activities. By analyzing user usage data and cross-matching them, libraries can examine the resource lists of the four foreign language databases, WoS publication statistics, WoS citation statistics, and denial access volumes from different dimensions. Figure 2 [Figure 2: see original paper] shows the types and quantities of documents published by Fudan University researchers in the four foreign language databases in 2020.

From Figure 2 [Figure 2: see original paper], we can see that among the four databases, Fudan University scholars published the most articles in Elsevier SD, followed by Wiley and Springer, and the fewest in Taylor-ST. In these four databases, the publication types are mostly articles, followed by reviews and meeting abstracts.

By calculating the annual download volume, cost per download, proportion of zero-download journals, citation volume, cost per citation, and download-to-citation ratio of each database, we can examine the annual utilization of the databases, as shown in Table 4 .

According to Table 4 , the cost per download and cost per citation vary across databases. For example, for the Taylor-ST database, the average cost per cited article is about 19,000 yuan. The download-to-citation ratio directly reflects the conversion rate from "information utilization" to "information creation." In the Springer database, for every 777 articles downloaded by users, one will be cited in publications, representing the highest conversion efficiency among the four databases.

3.3.2 Journals with Citations but No Holdings Since journals with citations but no holdings are also important basis for ordering strategy adjustments [35], libraries can take the year as a unit to count the databases where journals cited in the current year but not held, and where Fudan University scholars have published but are not held, to assist in ordering decision-making. By sorting the journals cited by Fudan University Library in 2020 but not held in descending order of citation frequency, and counting the top three most frequently cited journals, we find they are JAMA-JOURNAL OF THE AMERICAN MEDICAL ASSOCIATION, JOURNAL OF MATERIALS CHEMISTRY A, and BMJ-BRITISH MEDICAL JOURNAL. The databases where these journals are located are shown in Table 5 .

3.4 Construction of Usage Table Indicators

3.4.1 Coding Method This study uses grounded theory, a qualitative research method, to explore and construct indicator dimensions for the compre-

hensive usage table of readers' utilization of foreign language databases [36]. Grounded theory was jointly proposed by B. Glaser and A. Strauss in 1967, providing researchers with a complete set of methods and steps for inductively constructing theory from raw materials [37]. This study began with collecting and organizing the usage statistical data of four foreign language databases at Fudan University Library, which reveals the resource overview and reader utilization behavior of foreign language databases. Through coding, classification, and synthesis based on data characteristics, we obtained the connotation and relationships among conceptual categories.

Since the main goal of this study is to interpret the data itself, this paper adopts Strauss's strategy of conducting three types of coding simultaneously: open coding, axial coding, and selective coding.

(1) Open coding. In open coding, researchers are required to openly select data based on theoretical sensitivity, remain faithful to the original data during processing, and construct a brief and precise initial coding system. The usage of foreign language databases is reflected through many events, so this paper adopts an event-by-event approach for initial coding, obtaining 13 conceptual categories as the basis for axial coding. The open coding process is shown in Table 6 .

(2) Axial coding. Based on open coding, conceptual categories are summarized and relationships among categories are examined for axial coding. For example, in open coding, "annual download volume," "annual denial access volume," "cost per download," "proportion of zero-download journals," and "download-to-citation ratio" can all be summarized as "literature utilization." Finally, four axial coding concepts are formed.

(3) Selective coding. By systematically processing the relationships among conceptual categories obtained through axial coding, core and subcategories are further determined. It was found that the usage statistical data of the four foreign language databases at Fudan University Library can be mainly summarized into three dimensions: "resource guarantee," "supplementary measurement of academic achievements," and "bibliometrics." Resource guarantee includes "collection overview" and "subject guarantee"; supplementary measurement of academic achievements includes "literature utilization"; bibliometrics includes "literature feedback." The three-level coding results of this study are shown in Table 7 .

(4) Theoretical saturation test. This study uses reserved effective data from two analysis samples for theoretical saturation testing. The results show that the conceptual categories in the theoretical model have been fully developed, and no new dimensions were found beyond the three dimensions (resource guarantee, supplementary measurement of academic achievements, and bibliometrics) for analyzing the usage statistics of the four foreign language databases at Fudan University Library, nor were any new

initial concepts found internally. Therefore, it can be considered that the "Dimensions for Statistical Analysis of Four Foreign Language Database Usage at Fudan University Library" obtained through grounded theory has reached theoretical saturation.

3.4.2 Key Indicator Identification Based on the open coding in Table 6 and the utilization data of the four foreign language databases, we can obtain the utilization data of the four foreign language databases at Fudan University Library by year, as shown in Table 8 .

As shown in Table 8 , each indicator reflects the usage of the four foreign electronic journal databases ordered by the library from different perspectives. To further clarify the influence degree of each indicator in Table 8 and identify key indicators, this paper adopts the grey relational analysis method.

General abstract systems contain multiple factors whose combined effects determine the system's development trend. However, people often hope to identify the main factors among many—to know which factors have significant impact on system development and which have minor impact; which factors promote system development and need strengthening, and which hinder development and need suppression.

Regression analysis, variance analysis, and principal component analysis in mathematical statistics are all methods for system analysis. However, these methods have the following shortcomings [38]: First, they require large amounts of data support, otherwise statistical patterns are difficult to identify; second, they require samples to follow certain typical probability distributions, require linear relationships between factor data and system characteristic data, and require factors to be independent of each other—requirements that are often difficult to meet in reality; third, these methods may produce quantitative results that contradict qualitative analysis results, increasing the risk of distorting and reversing system relationships and patterns. Especially when statistical data is very limited, existing data has large grey levels, or many data points fluctuate dramatically due to human factors without typical distribution patterns, the above mathematical statistical methods are often ineffective.

Grey relational analysis compensates for these shortcomings. It has no specific requirements for sample size or patterns, involves small computational effort, is very convenient, and does not produce contradictions between quantitative and qualitative results [38]. The basic idea of grey relational analysis is to determine the connection between variables by analyzing the similarity of their geometric shapes. This method can clarify the "grey" relationships between factors and identify the main influencing factors of matters. It has been applied in many fields [39-41] and has reliability and broad applicability. This paper uses grey relational analysis to calculate Table 8 to clarify the influence degree of each indicator and identify key indicators. The specific identification steps are as follows:

(1) Determine the analysis sequences (parent sequence and sub-sequences). Take the maximum value of each column in Table 8 as the optimal value of the statistical results as the parent sequence $\{X_0\}$, obtaining $\{X_0\}$ = {2,645,774, 1,486,685, 581,885, 193,185}. The sub-sequences are total collection journals, net journal proportion, core journal proportion, "Double First-Class" subject journal matching rate, "Double First-Class" subject core journal guarantee rate, annual download volume, annual denial access volume, cost per download, proportion of zero-download journals, download-to-citation ratio, citation volume, cost per citation, and publication volume, denoted as X_1-X_{13} respectively.

(2) Preprocess variables to achieve dimensionless and reduce variable range for simplified calculation: first calculate the mean of each indicator, then divide each element in the indicator by its mean. The results are shown in Table 9 .

(3) Calculate the correlation coefficients between each indicator in the sub-sequences and the parent sequence. Let n be the sample size and m be the number of sub-sequences, then n = 4 and m = 13. For each indicator, calculate $|X_0(k) - X(k)|$, where $1 \leq i \leq 13$ and $1 \leq k \leq 4$, and denote $a = \min(i)\min(k)|X_0(k) - X(k)|$, $b = \max(i)\max(k)|X_0(k) - X(k)|$. The calculated matrix is shown in Table 10 .

(4) Calculate the grey relational degree V between each sub-sequence and the parent sequence according to the formula $V(X_0(k), x(k)) = (|X_0(k) - X(k)| + b) / (|X_0(k) - X(k)| + b)$, where is the resolution coefficient, generally 0.5. The overall relational degrees of each indicator are shown in Table 11 .

(5) Key indicator identification. The larger the value in the $?(X_0, X)$ column in Table 11 , the more important the indicator. Therefore, key evaluation indicators can be determined by sorting according to the size of $?(X_0, X)$. According to the calculation results in Table 11 , for Fudan University, the top five indicators affecting the usage effectiveness of the four foreign electronic journal databases (Elsevier SD, Wiley, Springer, and Taylor-ST) are annual download volume (1.0000), citation volume (0.9444), publication volume (0.8965), total collection journals (0.7442), and core journal proportion (0.7409).

3.5 Conclusions and Discussion

Research has shown that a grey relational degree value greater than 0.5 indicates that the comparison sequence has a relatively large influence on the reference sequence [42]. Through grey relational analysis of the constructed usage table of foreign electronic journal databases by Fudan University users, this paper calculated the influence degree of each indicator and identified key indicators. The results show that all overall relational degree values in Table 11 are higher than 0.5, indicating that the indicators are reasonably selected. The following

conclusions are drawn:

(1) In the supplementary measurement dimension, "annual download volume" has the largest relational degree value among all indicators, indicating that when considering foreign language database resource construction, libraries should first focus on this key indicator. Academic papers are the most direct and important research output of universities, and the degree to which literature resources in databases support university scholars' research is the main reason affecting university purchases. Researchers' full-text downloads after browsing databases are the most obvious first step for subsequent research and the most direct manifestation of reader database usage. Therefore, when evaluating purchased foreign language databases, libraries can use the database's annual download volume as a basic objective evaluation indicator, which is also consistent with previous library database selection experience. In addition, the grey relational degree values of both "annual denial access volume" and "proportion of zero-download journals" indicators exceed 0.7. Due to the complexity of research processes and user behaviors in using foreign language databases, annual download volumes of foreign language databases in different disciplines may show completely different situations, and even different foreign language databases under the same institution may have varying annual download volumes. Therefore, other important indicators need to be considered alongside annual download volume. Since denial access volume indicates unmet reader demand and zero-download journals indicate resources that readers may not need, both are also important basis for ordering strategy adjustments.

(2) In the bibliometrics dimension, the grey relational degree value of the "citation volume" indicator exceeds 0.9, and the "publication volume" indicator reaches 0.8965, indicating that these two indicators are extremely important in foreign language database resource construction in university libraries. After literature materials in databases are downloaded by users, if they are utilized in subsequent research, the most direct manifestation is citation in research results, which also shows that the database's content materials promote the generation of innovative thinking among university researchers. For senior scholars, the more papers they publish, the higher the probability of publishing high-impact papers in their discipline [43]. Since academic papers are the main research output of universities, the publication volume of university scholars in a certain foreign language database will influence research by university researchers. Therefore, as a literature information center supporting faculty and student research, university libraries must also consider the inclusion of the foreign language database in the university's research publications when conducting resource construction. At the same time, although the second and third highest grey relational degree values (citation volume and publication volume) both belong to the bibliometrics dimension, the "cost per citation" indicator has the lowest grey relational degree value. Therefore, when

evaluating user usage and formulating foreign language database ordering strategies, other indicators can be prioritized.

(3) In the resource guarantee dimension, the grey relational degree values of both "total collection journals" and "core journal proportion" indicators are higher than 0.74, indicating that both the overall collection volume and quality should be considered in foreign language database construction. Total collection journals represent the breadth of content coverage of the foreign language database used by the university, while "core journal proportion" further reveals the quality depth of the foreign language database's content. Both are indispensable for supporting university scholars' research. Therefore, when planning collection resource construction, university libraries need to consider both the breadth and depth of foreign language database content according to the university's own discipline construction planning. In recent years, with the proposal and advancement of the national "Double First-Class" initiative, the principle of "adhering to discipline-based construction" has also become an important guiding ideology for university library resource construction [44]. The calculation results also show that the grey relational degree values of "Double First-Class" subject core journal guarantee rate and "Double First-Class" subject journal matching rate are as high as 0.7343 and 0.6626, indicating that to build first-class disciplines, universities need matching, high-quality (core) literature resources as basic conditions and important support.

4 Implications for Foreign Language Database Construction in University Libraries

4.1 Data-Driven Analysis of Collection Resources

The "Recommendations of the CPC Central Committee on Formulating the 14th Five-Year Plan for National Economic and Social Development and the Long-Range Objectives Through the Year 2035" [45] proposes "promoting the digitalization of public culture." As public literature information resource centers within universities, strengthening information resource construction is an eternal goal of libraries, requiring them to build information resource systems considering different resource characteristics and reader usage under current development contexts. In the digital transformation of collection resources, for the selection and collection construction of foreign language databases, libraries should be service-centered for faculty and students, aiming to support university discipline construction, especially the "Double First-Class" discipline construction. Therefore, if data-driven analysis of collection construction status and reader utilization can be conducted, libraries can understand the distribution, quality, guarantee for various disciplines, and utilization of collection resources. Based on the usage table of foreign language databases and by calculating key targets that should be emphasized in database resource construction, libraries

can reduce obstacles in structural setting, selection, evaluation, and other aspects of foreign language database resource construction, avoid the dilemma of "treating everything equally" caused by language differences in foreign resource management, and strengthen the theoretical basis and intelligence level of collection resource management.

4.2 Mining the Research Value of Download Volume, Citation Volume, Publication Volume, and Related Data

Foreign language databases retain "traces" of user utilization. Through grey relational analysis, this paper finds that in constructing foreign language database resources, libraries should not only pay attention to the total collection journals and proportion of core journals purchased but also strengthen the analysis and research of annual download volume, citation volume, publication volume, and related data, leveraging their research value in academic evaluation, information retrieval, scientific knowledge 梳理 and visualization, and discipline planning and development. Regarding download volume and related data, foreign academic literature downloading behavior is an important part of researchers' literature retrieval behavior. University libraries' attention to download volume is significant for strengthening foreign language database construction. First, predicting users' literature download volume from collection foreign language databases helps deeply understand users' retrieval behavior in foreign language databases, provides basis for optimizing retrieval results and reconstructing ranking of the university's academic resource retrieval platform, thereby improving retrieval system service quality [46]. Second, exploring the correlation between download volume and other factors, such as the correlation between download volume and publication volume or citation frequency, helps libraries determine performance evaluation indicators for foreign language databases and enhance the scientific nature of library decision-making. Finally, focusing on download volume and user preference research can not only visualize user usage of foreign language databases but also help libraries gain deeper scientific understanding of user behavior in utilizing foreign language databases, thereby creating more accurate user portraits and providing personalized services. Regarding citation volume and related data, starting from foreign language database citation volume and conducting mining research on cited literature's citation frequency, citation location, and citation text themes can enable descriptive statistics of foreign language databases and help construct citation networks. Regarding publication volume and related data, libraries can use various visualization tools for multidimensional quantitative and research topic clustering analysis of certain fields, revealing development trends, research strengths, and hotspots in those fields, and conducting foreign language database construction from the perspective of discipline planning and development.

4.3 Digital Resource Sharing and Co-construction

Currently, foreign language database construction in Chinese university libraries faces issues such as funding pressure and resource structure [13]. Resource sharing and co-construction among libraries can improve the integration and perfection of information resources, representing a further requirement for foreign language database resource construction in university libraries in the intelligent era. When formulating database resource construction strategies, university libraries should consider not only their own purchased foreign language databases but also resource sharing and co-construction with other university libraries, public libraries, and government enterprises to deepen cooperation. First, university libraries can participate in university library consortia to build a unified foreign language data resource platform for data sharing and unified development and utilization among consortium members, providing effective guarantees for data resource utilization and analysis of member libraries. They should also consider the characteristics of different discipline backgrounds regarding reader resource types and usage habits, with emphasis in resource co-construction cooperation. Second, they should actively explore unique and characteristic literature resources. By leveraging the advantages of key university disciplines, conducting in-depth analysis of existing foreign language collection resources, and collecting and organizing these resources, they can build a characteristic resource system for self-use and sharing. Third, university libraries must also pay attention to privacy protection and ownership protection in sharing, adopting supporting measures such as strengthening library privacy protection supervision systems, establishing privacy protection supervision systems in scientific data sharing, and establishing privacy risk assessment mechanisms [47] to avoid unnecessary resource rights disputes. Finally, other university libraries can also follow the 思路 and methods of this paper to conduct real-time analysis and control of their foreign language resource co-construction from a data perspective.

This study, based on user information behavior theory and combined with existing research and actual characteristics of university collection resource construction, organizes and summarizes specific data on foreign language database usage from different perspectives, analyzes the 2020 user usage of four foreign language databases at Fudan University Library, uses grounded theory to code and summarize a comprehensive usage table of foreign language databases, and finally uses grey relational analysis to calculate and evaluate the influence degree of each indicator, identify key indicators, and provide strategic suggestions for foreign language database construction in university libraries. The characteristics and contributions of this paper lie in: Expanding the analytical perspective of library resource construction based on the context of user utilization of foreign language databases, with empirical analysis using Fudan University Library as an example to further confirm the practicality and applicability of the method.

Using grounded theory to code, classify, and synthesize statistical data on user utilization of foreign language databases to construct a usage table of foreign language databases by Fudan University users as the basis for subsequent

analysis. Successfully identifying key indicators using grey relational analysis, providing methodological references and key targets, basis, and ideas for subsequent research on information resource construction in universities.

This study also has some limitations: First, indicators can be further refined, such as subject classification can be detailed to first-level and second-level disciplines, and denial access volume can be calculated specifically under each subject category. Second, the sample size used in the study is small, only including four foreign electronic journal databases. Future research can include more databases (such as open access databases) to expand the sample scale.

References

[1] Higher Education Library and Information Work Guidance Committee of Ministry of Education. 2019 Statistics on Electronic Resource Procurement Expenses of 1,165 University Libraries [EB/OL]. [2021-07-30]. http://www.scal.edu.cn/tjpg/202011160231.

[2] Higher Education Library and Information Work Guidance Committee of Ministry of Education. 2019 Statistics on Literature Resource Procurement Expenses of 1,263 University Libraries [EB/OL]. [2021-07-30]. http://www.scal.edu.cn/tjpg/202011160231.

[3] Higher Education Library and Information Work Guidance Committee of Ministry of Education. 2019 Statistics on Annual Total Funding of 1,278 University Libraries [EB/OL]. [2021-07-30]. http://www.scal.edu.cn/tjpg/202011160231.

[4] Chinese University of Hong Kong Library. 2019-2020 Annual Report [EB/OL]. [2021-07-30]. https://www.lib.cuhk.edu.hk/sites/cuhk/files/page/about/inside_{{{the}}}*{{library}}* 20.pdf.

[5] Tsinghua University Library. Tsinghua University Database Navigation System - xserver version [EB/OL]. [2021-07-28]. http://nav.lib.tsinghua.edu.cn/xport/dbdh.htm.

[6] Peking University Library. Database Navigation [EB/OL]. [2021-07-28]. https://dbnav.lib.pku.edu.cn/.

[7] Fudan University Library. Academic Resources Portal [EB/OL]. [2021-07-28]. http://202.120.227.56:8331/V/24N8BV6BNHSTTGCD4IDD74LELC4915ANHB42775HC9H4ALFX1-01000?FUNC=FIND-DB-1-LCL&pds_{handle}=GUEST.

[8] Shanghai Jiao Tong University Library. Database List [EB/OL]. [2021-07-28]. http://www.lib.sjtu.edu.cn/f/database/database.shtml.

[9] University of Hong Kong Libraries. Databases [EB/OL]. [2021-10-30]. https://libguides.lib.hku.hk/az.php.

[10] Hong Kong University of Science and Technology Library. All databases - HKUST Library [EB/OL]. [2021-10-30]. https://library.ust.hk/collections-resources/databases/all-databases/.

[11] Chinese University of Hong Kong Library. A-Z Databases [EB/OL]. [2021-10-30]. https://libguides.lib.cuhk.edu.hk/az.php.

[12] Zhejiang University Library. Database Navigation [EB/OL]. [2021-10-30]. http://210.32.137.90/s/lib/libtb/.

[13] Su Jianhua. Dilemmas and Countermeasures in Foreign Electronic Journal Procurement in Domestic University Libraries [J]. New Century Library, 2016(7): 48-50.

[14] Wang Chunsheng. "Cheng Huanwen's Question": Causes and Solutions—Reflections on Foreign Journal Database Procurement [J]. Library Forum, 2015, 35(5): 1-5.

[15] Zhao Yanzhi. Strategies of Foreign Libraries in Dealing with the Crisis of Big Deals and Their Implications for China [J]. Library Work and Research, 2015(12): 61-64.

[16] Fang Shiyu, Hao Qun, Zhang Libin. Investigation and Analysis of Licensing Agreement Models Used in Digital Resource Procurement by American University Libraries [J]. Information Studies: Theory & Application, 2021, 44(8): 187-195, 158.

[17] Hu Sai. Transaction Dilemmas and Coping Strategies of Electronic Journals in University Libraries—From the Perspective of Data Vendors [J]. University Library Work, 2018, 38(5): 45-49.

[18] Xiang Confucius, Qin Xia. Performance Evaluation of Foreign Electronic Journal Full-text Database Usage in University Libraries—Taking South China University of Technology Library as an Example [J]. Information Research, 2020(10): 95-101.

[19] Bi Yanfang, Li Taifeng. Digital Resource Evaluation Method and Example from User Data Dimension—Taking University of Electronic Science and Technology Library as an Example [J]. Library and Information Service, 2017, 61(22): 82-88.

[20] Tang Ganghui, Cheng Huanwen. Analysis of Wiley Journal Database Usage Statistical Report—Taking Sun Yat-sen University Library as an Example [J]. Library and Information Service, 2007, 51(4): 136-138, 149.

[21] Wang Qiong, Wu Yu, Wu Yingmei. Research on Foreign Academic Library Service Capability Evaluation System [J]. Journal of Academic Libraries, 2009, 27(2): 85-90.

[22] Xiang Yuanyuan, Wen Guoqiang. Exploration on University Library Effectiveness Evaluation and Its Indicator System Construction [J]. Journal of Academic Libraries, 2011, 29(3): 11-16.

[23] Shanghai Jiao Tong University Library Annual Report. Content Details [EB/OL]. [2021-07-28]. http://www.lib.sjtu.edu.cn/f/content/detail.shtml;jsessionid=30cb4967-37eb-40f9-b5e2-8b464f6b9d43?id=6302&lang=zh-cn.

[24] Yang Xin, Zhang Min, Liao Jianlan, et al. Research on Electronic Resource Utilization from the Perspective of Social Network—Taking Fudan University Library as an Example [J]. Library and Information Service, 2021, 65(15): 91-99.

[25] UCLibrarySearch. UC Libraries [EB/OL]. [2021-07-30]. https://libraries.universityofcalifornia.edu/uclibra

[26] TU Delft and the National Library of the Netherlands launch Future Libraries Lab [EB/OL]. [2021-07-30]. https://www.tudelft.nl/en/2020/tu-delft/tu-delft-and-the-national-library-of-the-netherlands-launch-future-libraries-lab.

[27] University of Oxford GLAM Digital Programme. Gardens, libraries & museums [EB/OL]. [2021-07-30]. https://www.glam.ox.ac.uk/digital-strategy.

[28] Wang Wei, Wang Lin. Research on the Correlation Between Foreign Electronic Journal Utilization Behavior and Research Output in University Libraries [J]. Information Research, 2016(12): 11-16.

[29] Liu Jiayin. Research on Comprehensive Evaluation of Digital Resource Utilization Efficiency in Domestic University Libraries [D]. Wuhan: Wuhan University, 2013.

[30] Boukacem-Zeghmouri C, Bador P, Lafouge T, et al. Relationships between consumption, publication and impact in French universities in a value perspective: a bibliometric analysis [J]. Scientometrics, 2016, 106: 263-280.

[31] Nicholas D, Rowlands I, Huntington P, et al. E-journals: their use, value and impact [EB/OL]. [2021-06-30]. https://www.researchgate.net/publication/266034523_E-journals_{{{their}}}*{{use}}*{{{value}}}*{{and}}*{impact}.

[32] Jung Y, Kim J, So M, et al. Statistical relationships between journal use and research output at academic institutions in South Korea [J]. Scientometrics, 2015, 103: 751-777.

[33] Liu Jiayin. Statistical Analysis of Electronic Resource Usage and User Retrieval Behavior in University Libraries—Taking ScienceDirect Database as an Example [J]. Journal of Academic Libraries, 2012, 30(2): 81-86.

[34] Zhuang Jilin, Chen Ling, Yao Xiaoxia. Quantitative Research on Content Quality Indicators of Foreign Journal Databases [J]. Library and Information Service, 2021, 65(2): 54-64.

[35] Li Na. Electronic Resource Management and Evaluation—Taking Foreign Electronic Journals of Shanghai University of Finance and Economics Library as an Example [J]. Library and Information Service, 2018, 62(15): 60-62.

[36] Mi Ziyuan, Zhang Libin, Ding Kening. Analysis of Improving Electronic Book Utilization Efficiency in University Libraries Based on Grounded Theory [J]. Information Science, 2019, 37(7): 48-55.

[37] Ke Ping, Zhang Wenliang, Li Xining, et al. Research on Librarians' Perception of Organizational Culture in Public Libraries Based on Grounded Theory

[J]. Journal of Library Science in China, 2014, 40(3): 37-63.

[38] Liu Sifeng, Dang Yaoguo, et al. Grey System Theory and Its Application (5th Edition) [M]. Beijing: Science Press, 2010.

[39] Zhang Xinxiang, Duan Yanhong. MOOC Quality Judgment Based on Learner Online Review Text—Taking Online Review Text from "Chinese University MOOC" as an Example [J]. Modern Educational Technology, 2020, 30(9): 56-63.

[40] Zhang Kun, Wang Xue, Li Li. Construction and Analysis of Service Quality Evaluation Indicator System for University Library WeChat Official Accounts [J/OL]. Library Construction: 1-11 [2021-08-11]. http://kns.cnki.net/kcms/detail/23.1331.G2.20210423.1455.008.html.

[41] Niu Yongping. Grey Relational Analysis on Efficiency Influencing Factors of Public Libraries [J]. Library Journal, 2012, 31(2): 5-10.

[42] Pu Xiaoge, Liu Liming. Coefficient of Variation and Grey Relational Analysis of Database Performance and Empirical Research [J]. Library and Information Service, 2014, 58(14): 71-78.

[43] Lariviere V, Costas R. How many is too many? On the relationship between research productivity and impact [J]. PLoS One, 2016, 11(9): e0162709.

[44] Zhang Yue, Ke Jiaxiu, Du Li. Method and Empirical Analysis of Priority Guarantee for University Discipline Journal Resources [J]. Library and Information Service, 2020, 64(8): 71-77.

[45] Recommendations of the CPC Central Committee on Formulating the 14th Five-Year Plan for National Economic and Social Development and the Long-Range Objectives Through the Year 2035 [EB/OL]. [2021-01-25]. http://china.cnr.cn/news/20201104/t20201104_{525318774}.shtml.

[46] Xie Hao, Wu Xuehua, Chen Qian, et al. Research on Predicting Academic Literature Download Behavior by Integrating Multi-dimensional Features [J]. Library and Information Service, 2021, 65(12): 112-121.

[47] Zhu Bei, Wang Chuanqing. Analysis of Privacy Governance in Scientific Data Sharing in China [J]. Library and Information Service, 2020, 64(22): 37-47.

Author Contributions:

Zhao Weijia: Paper writing;
Hao Qun: Topic design, writing guidance, partial content writing;
Zhang Libin: System architecture, partial content writing, paper revision.

The Usage Statistics and Analytical Thinking of Foreign Language Database Resource Construction in University Libraries: Taking the

Practice of Four Databases in Fudan University as an Example

Zhao Weijia[1], Hao Qun[1], Zhang Libin[2]
[1]Fudan University Library, Shanghai 200433
[2]Nankai University Library, Tianjin 300071

Abstract: [Purpose/significance] Since the 21st century, the proportion of databases in the procurement resources of university libraries has been continuously increasing, and foreign language databases have become an important source of electronic resources. At the same time, in the face of the strict requirements of various supervision and audit work, staff have also encountered many problems in the process of purchasing and constructing foreign language database resources. Taking Fudan University Library as an example, through the analysis and research on the resource ordering and usage of four foreign language databases (Elsevier SD, Wiley, Springer, and Taylor-ST), this paper provides ideas and references for the construction decision-making of foreign language databases in university libraries. [Method/process] According to the statistical data of the four foreign language databases, the general situation of collection resources, the guarantee situation of "double first-class" discipline construction, the annual utilization situation, and the situation of annual highly-cited journals without holdings are classified and summarized. From these four analysis perspectives, the data is coded based on grounded theory, and a comprehensive usage table of foreign language databases is obtained based on specific data, and the influence degree of each index is calculated and evaluated by using grey relational analysis method. [Result/conclusion] Through calculation, five key indicators are identified that most affect the use of foreign language databases in Fudan University Library: annual downloads, citations, publications, total number of journals in the collection, and the proportion of core journals. From three perspectives—supplementary measurement of academic achievements, bibliometrics, and resource guarantee—optimization suggestions are proposed for the construction of foreign language databases in Chinese university libraries, providing methodological references and key targets, basis, and ideas for subsequent research on information resource construction in other universities.

Keywords: university library; foreign language database; resource construction; grounded theory; grey relational analysis

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