

Application and Development Trend of Data Visualization in Library

Fenghua Lv^{1,*}, Jianing Ma², Xiaomei Zhang¹

¹Qingdao University Library, Qingdao, Shandong, China.

²School of Rehabilitation, Shandong University of Traditional Chinese Medicine, Jinan, Shandong, China.

*Corresponding Author.

Abstract: For illustrating the application practice of data visualization in library, and taking the journal papers collected in CNKI journal database as the research object, Excel 2016 data statistical software and VOSviewer1.6.19 and other visual analysis software are used. The research results of the application of data visualization in library are divided into four stages for visual analysis of research hotspots and trends. The research focus of data visualization application in library industry has gradually shifted from data analysis to data service. With the widespread use of social media, libraries pay more attention than ever to the cultivation of user data literacy, the digitization of knowledge services, the visualization of information at a glance, the research of digital humanities. Data visualization technology opens a window for users to gain insight into library knowledge resources with data, and speeds up the process of transformation from digital library to smart library.

Keywords: Data Visualization; the Library; VOSviewer; Knowledge Service; Intelligent Library

1. Introduction

Vision plays an absolutely dominant role in the process of human perception of external information. The advantage of visual analysis is that it can be clear at a glance and play the effect of "one picture shows spring and autumn". Data visualization technology has strong application potential in library industry. The innovative application of data visualization means and methods combined with library characteristics can improve the level of library management. The library industry also pays more and more attention to

the application of visualization technology.

2. Concept and Characteristics of Data Visualization

2.1 Concepts

Data visualization is the use of data analysis technology, data visualization tools, in the form of graphics, images, to visualize the special data in the mass data to discover the information hidden in the mass data [1]. Data acquisition, preprocessing, storage and mining have a great impact on data visualization.

2.2 Features

2.2.1 Intuitive and Clear

The processing speed of human brain for visual information is much faster than that of text information. Data visualization enables people to visually insight into the relationship characteristics and rules behind massive data, so as to facilitate rapid understanding of the acquired information and make corresponding decisions. Data visualization has both interactivity and excellent visualization effect, and the display effect is better than the traditional statistical chart, which improves the user experience [2]. Through the use of different expression methods, the information elements are designed, and the abstract information and data are transmitted by intuitive expression techniques through graphics, charts, colors and words, and the complex information is simplified, visualized and organized.

2.2.2 Quickly Reflect the Status of Data Updates

By updating data information in real time, data visualization technology facilitates library managers to quickly identify new opportunities or problems and make rectification decisions in time.

2.2.3 Facilitate Smooth Communication

Data visualization eliminates interference information, avoids vague language and ineffective communication, and clearly presents key content, improving communication efficiency.

3. Application Practice of Data Visualization in Library Industry

Data visualization technology can be used in library management and service. It can not only provide bibliographic data and subject research assistance for readers, but also be used to evaluate the library and improve the collection structure.

3.1 Knowledge Service Management

3.1.1 Visualization of Paper Book Lending

Paper document borrowing is the basic function of library knowledge service. The visualization technology can show the borrowing situation of readers at a glance according to different criteria such as subject, time period and borrowing amount. Since the borrowing situation can be updated in real time, it is easy to find changes at any time and take corresponding measures. Based on big data, the user portrait visualization framework is established, the utilization rate of library resources is studied by combining the borrowing behavior data of users, and the reading preference of readers is related to the optimization of library structure.

3.1.2 Facilitate the Retrieval and Analysis of Digital Collections

Today's libraries are in the process of transforming from digital to intelligent. The library can digitize the digital collection (such as the representative catalogue of ancient books of the past dynasties), and carry out visual analysis on this basis. By building an interactive visual analysis system, to provide an overall and effective research tool for digital collection document analysis based on large-scale data [3]. In order to realize subject-based visualization of search results in digital library, Bakhshayesh et al proposed subject-based document retrieval. The document is used as the carrier of the topic, and then the vector space model is used for information retrieval, retrieving the documents related to the user's query, while visualizing each document through a ring graph, displaying the topic in each document and

query. Bakhshayesh et al. also evaluated their proposed topic-based retrieval and visualization methods from multiple perspectives, the results: 67% of users prefer topic-based document retrieval, and 80% of them consider the proposed visualization method to be practical [4].

3.1.3 Discipline Service

With the help of data visualization, users can be provided with a variety of disciplinary services. Take subject analysis service and benchmarking service as an example. Subject Analysis services can provide subject based analysis reports for school research departments. For example, ESI analysis, through the regular detection of the overall level of scientific research of the entire university, can have an objective understanding of the overall scientific research performance of the university, the status quo of dominant disciplines, and the development trend. Then the potential subject mining and development trend forecast. If a university wants to know which disciplines are more likely to enter the top 1% of ESI in order to focus on the process of future research support, it can first calculate the total citation frequency (i.e., the threshold value) of the last institution to enter the top 1% of the discipline in each ESI discipline. The total citation frequency of each discipline in the university is compared with the threshold value of the top 1% in the world for each discipline.

For example, suppose the comparative analysis chart of the total cited frequency and threshold value calculated by various disciplines in a university is shown in Figure 1.

As can be seen from Figure 1, Subject 1 has entered the top 1% of ESI globally, subject 2, Subject 3 and subject 4 belong to the strong disciplines of the university, but in comparison, subject 2 is the closest to the top 1% in the world. If the university gives more scientific research support, it is likely to accelerate the threshold of entering the top 1% in the world. Benchmarking analysis can be carried out on the subject fields of relevant universities at home or abroad to understand the research status and development trends of different research fields and high-level research teams. China National Knowledge Network (CNKI) was used as the research literature source database, the search scope was academic journals, the search method was $SU = ("$

rehabilitation "and" Traditional Chinese Medicine"), the research literature was published as of May 31, 2023, and the author units were Shandong University of Traditional Chinese Medicine and Heilongjiang University of Traditional Chinese Medicine. Thus, the research situation of the two Chinese medicine universities was benchmarked and analyzed, as shown in Figure 2.

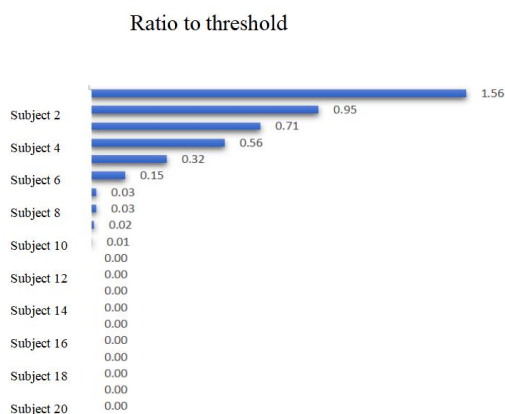


Figure 1. Comparative analysis of the total citation frequency and threshold value calculated by various disciplines in a university

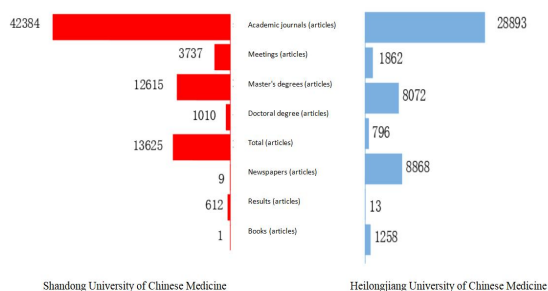


Figure 2. Comparison of scientific research results published by Shandong University of Chinese Medicine and Heilongjiang University of Chinese Medicine in the field of TCM rehabilitation

From the comparison, we can make a preliminary judgment that in the field of TCM rehabilitation research, the total number of master and doctoral papers of Shandong University of Chinese Medicine far exceeds that of Heilongjiang University of Chinese Medicine, indicating that in terms of the scale of running a school in this field, the scale of running a school of Shan Chinese Medicine is larger than that of Black Chinese medicine, and the number of academic journal papers and conference papers published by Shan Chinese medicine in the field of TCM

rehabilitation exceeds that of black Chinese medicine. Researchers in the field of TCM rehabilitation are willing to participate in more conferences and communicate with their peers, and have published more conference papers. The number of achievements of black Chinese medicine is nearly twice that of mountain Chinese medicine, indicating that the scientific research work of black Chinese medicine is more focused on the results, and the strength of this aspect is stronger.

3.1.4 Academic Guidance

Through the analysis of research hotspots and the status of research teams, academic guidance is provided to researchers. For example, software such as VOSviewer can be used to grasp research hotspots, development trends and future directions by analyzing the time distribution, journal distribution and subject distribution of relevant literature and conducting keyword clustering and timely area analysis. In order to provide a better academic development environment in the aspects of policy guidance, scientific research funding, and academic exchanges.

To understand research trends in a field, you can use trend charts. For example, the trend chart made by taking TCM rehabilitation as the subject term and the year and number of journal papers as variables is shown in Figure 3.

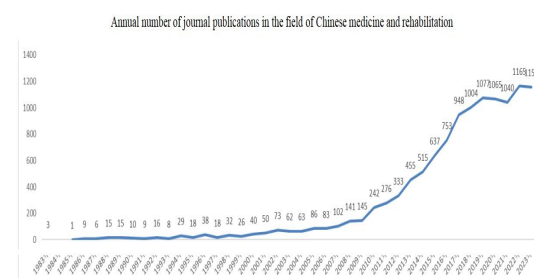


Figure 3. The number of journals published in the field of TCM rehabilitation

As can be seen from Figure 3, the number of papers in the field of TCM rehabilitation has increased rapidly since 2010, indicating that after 2010, the research of TCM in clinical rehabilitation treatment has become more and more in-depth, the treatment means have become more diversified, the application has become more and more extensive, and the research results have shown a rising trend.

Through the visual quantitative analysis of the literature, we can also analyze the research hotspots of TCM rehabilitation therapy. It can

be seen that the top ten key research areas of TCM rehabilitation are stroke, clinical research, clinical observation, TCM nursing, rehabilitation therapy, TCM rehabilitation nursing, post-stroke rehabilitation training and clinical efficacy observation, as shown in Figure 4.

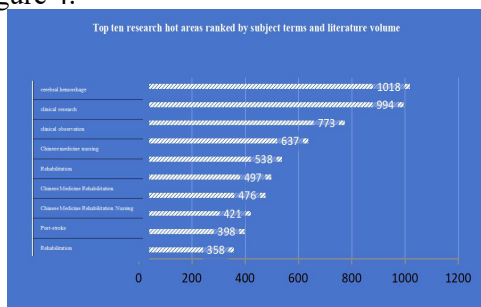


Figure 4. The top ten hot research areas of TCM rehabilitation therapy

3.2 Human Resource Management

The use of charts and other visual methods can not only visually display the scores of various abilities of librarians, but also display and compare the workload and scientific research contributions of librarians at a glance by setting different variables. If the business ability, research ability, work attitude, communication skills and ability to learn five dimensions for A department five librarians' scores, the results of A (90,80,95,75,85) for librarians, librarians B (85,70,85,85,80), librarians C,90,85,85,90 (85), Librarian D (65,40,80,85,60), Librarian E (85,60,90,90,80). From words and figures alone, it is difficult to see the difference at a glance, and repeated comparisons are needed. However, using the radar map, the difference in ability between the long edition and the short board of each librarian can be clearly seen, as shown in Figure 5.

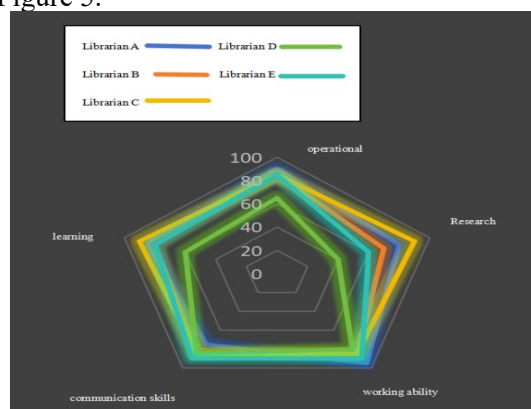


Figure 5. Radar chart of five librarians in a department

In addition, to evaluate the workload of librarians, slicers can be used to make the result statistics of the workload and scientific research volume of individuals in different years and departments, carry out phased monitoring and statistical analysis of the workload or scientific research output of librarians, and visually display the workload or academic output.

4. The Application of Large Visual Screens in Libraries

Data visualization large screen display is a combination of visual communication, computer technology and user psychology, using the large screen as a display carrier for intuitive, cool and scientific and technological display. After sorting out the business indicators to determine the requirements, the defined indicator information is mapped into visual elements, and delivered to the viewer through touch, interactive operation or one-way information display [5].

Visual charts are extremely powerful. The bar chart is used to show the same type of data over time, and the library visualization screen can use the bar chart to quickly discover which books are checked out in the library and which books are more interesting to readers. Line charts can be used to reflect the change and development trend of data over time. Through the line chart, the number of people entering the museum in different time periods can be found, which is convenient to adjust the peak flow of people in the peak period. Study room reservation can be used to clearly show the relationship between the whole and the part, and clearly show the efficiency of the use of self-study seats.

5. Challenges Faced by Data Visualization in Library Applications

5.1 Introduce and Train New Librarians to Meet Talent Challenges

To do a good job of data visualization in library, librarians need not only to know the business of library, but also to have data thinking and data visualization skills. The processing of multi-source data of different types and structures also poses challenges to the professional competence of librarians. The library generates a large number of borrowing data, user data, click volume and other data,

and the scale of the data is expanding over time, and the data changes quickly and is mostly traffic data. It is not only necessary to maintain the processing capacity of single machine, external memory model and even small computing cluster, but also to pay attention to the quality of the data in the process of data acquisition and analysis. Reasonable control of data uncertainty. The requirements for librarians' ability of real-time data analysis and data visualization are also high. Intelligent library requires librarians to have data thinking, data analysis technology and high information service literacy. The traditional allocation of library human resources can not meet the requirements of this new intelligent library for the quality of librarians. It is necessary to pay attention to the training of data visualization service awareness and skills of in-service librarians while introducing relevant talents.

5.2 Ensure that the Data base of Visual Information Services is Accurate and Secure

As an important basis for the application of visualization technology, it is necessary to ensure accurate data collection, timely upload the collected data to the overall system, and pay attention to information security in the process of data collection to prevent data loss, identity information and other privacy information from being stolen in the system. In the process of data visualization design preprocessing, data mining technology is used to statute, extract, clean, integrate and transform the data. Through data preprocessing, appropriate data sources are selected and target data is screened. At the same time, it is necessary to verify the data and check whether the data is standardized and whether there is data conflict [6].

5.3 Innovative Services in the Transformation from Digital Library to Smart Library

Data visualization includes animated interactive views, multi-dimensional drawings, maps, etc. Data visualization technology provides a powerful means for the library to do a good job in knowledge service. Ivy University libraries in the United States established different characteristics of data visualization space to enrich the means and

ways of visual expression [7]. In order to realize the transformation into a smart library and provide users with better knowledge services, libraries need to develop more innovative services, think about the operation of libraries from a more global perspective, and help to cooperate with internal librarians [8]. To better provide users with data visualization services requires multi-department cooperation in libraries, which helps to collect and share data from a global perspective. Through data visualization, information is presented in a way that is better for the brain to spot patterns or trends [9]. The user-centered concept should be integrated in the entire operation process of the library, always adhere to the user orientation, and the setting of data connection should be accurate and appropriate, simple and flexible. The operation interface is easy to understand, avoiding over-connection and invalid connection. Through the visual display of data, it is convenient for library managers to perceive, capture, record, mine data, analyze readers' needs and other means to provide more intelligent services [10].

6. Research Hotspots and Future Development of Data Visualization in Library Industry

Visual analysis of scientific research literature can be used to find research hotspots and development trends. This paper takes the research hotspot of data visualization in library industry as an example to find the research trend of data visualization application in library industry.

6.1 Data Sources

China National Knowledge Network (CNKI) is used as the research literature source database, the search scope is academic journals, and the search formula is $SU = (\text{"data visualization"} \text{and} \text{"library"})$.

6.2 Research Methods and Processes

The data deadline is set to June 15, 2023, and the advanced search is carried out with data visualization and library as the subject words, and 174 journal papers are obtained. Through the analysis of the time distribution of relevant literature and the cluster analysis of keywords, this paper divides into four stages to study the application of data visualization in the field of academic journal papers and research hotspots

and future development trends.

A search in the CNKI database shows that the earliest reference to data visualization was published in 2000. From January 1, 2000 to June 16, 2008, there were only 5 relevant journal papers, indicating that the application of data visualization in library research was still in its infancy during this period. The research content belongs to the interdisciplinary research of computer science and library and information science, and the research is conducted in a five-year period from June 2008 to June 2023. According to the time, it is divided into the initial development stage (2008-2013) (11 articles), the development stage (2013-2018) (67 articles) and the prosperity stage: In the last five years, 2018-2023 (91 articles).

The title, author, source and key words of related literature are exported in Endnote format as sample data, and visualization analysis is performed using Vosviewer1.6.19 software. Select Co-occurrence for Type of analysis, Keywords for Unit of analysis, full counting for calculation method, 2 for Attraction and 0 for repulsion. Set resolution in clustering to 1.00 and min. cluster size to 1. Term is set to 1 and 2 respectively for the germination stage and the preliminary development stage due to the small number of journal papers. Term is set to 3 for the development stage and prosperity stage due to the large number of journal papers [11].

6.3 Result Analysis

From the point of view of the number of published papers, from the germination to the prosperity stage, the number of published papers increased rapidly, indicating that the research content is more and more extensive, and the results are more and more rich.

The visualization results of hot data of journal paper research conducted in four stages from 2000 to 2023 are shown in Figure 6.

The comparison of research hotspots in four stages of the application of data visualization technology in library after data cleaning is shown in Table 1.

After the software analysis results are cleaned and the words such as library and data visualization are removed for comparison, it can be found that the research on the application of data visualization in the field of library is more and more focused on the

service goal of smart library [12].

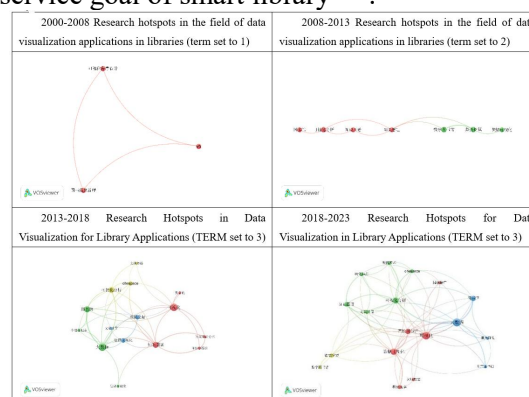


Figure 6. Research hotspots of data visualization in different stages of library application

Embryonic stage: 2000-2008.

Because there are only five literatures, even if the term is set to 1, the five articles only focus on three hot research areas: vb(visualization), visual programming and book information management. As a cross-discipline of library science and computer science, the application of data visualization is mainly focused on visual programming, library information management, such as the description and classification of library collections, and the use of visualization to display the category distribution of collections or the number of books on a specific topic [13]. Use Excel and other software to make simple statistics and analysis of user access statistics and borrowing data.

Initial development stage: 2008-2013.

In the study of 11 literatures from 2008 to 2013, if term is set to 3, only three co-occurrences of digital library, digital mining and visual analysis can be obtained; if term is set to 1, 37 words can be obtained in scattered distribution, and it is difficult to see hot spots, so term is set to 2. After the two words library and data visualization are removed, the five co-occurrence words data mining, visual analysis, knowledge management, digital library and knowledge graph can be obtained. At this initial stage of development, it can be seen that the research has advanced to the research direction of digital library through the use of data mining and data visualization analysis. During this period, libraries began to use more advanced data visualization tools for more complex data analysis [14]. It also uses visual data analysis to reveal user behavior and borrowing patterns to

help libraries better understand user needs and preferences.

Table 1. Comparison of High-frequency Words in Different Research Stages of the Application of Data Visualization Technology in Libraries

2000-2008 (5 piece)		2008-2013 (11 piece)		2013-2018 (67 piece)		2018-2023 (91 piece)	
high frequency word	Frequency	high frequency word	Frequency	high frequency word	Frequency	high frequency word	Frequency
vb	1	data mining	3	big data	21	big data	27
visual programming	1	visual analysis	3	knowledge map +citespace	14 +7	knowledge map	9
library information management	1	digital library	3	data mining	8	data mining	7
		knowledge map	2	Personalized service	3	Bibliometrics	6
		knowledge management	2	content analysis method	3	data service	6
				statistical analysis of word frequency	3	Smart Library	5
				digitize	3	Data literacy	5
				Literature Transfer	3	Reading Promotion	4
				Bibliometrics	3	digital library	4
				Information visualization	3	Linked Data	4
						digital humanities	3

Development stage: 2013-2018.

The application of more and more rich visual software tools began to prevail. Citespace software and various knowledge graphs have been frequently used in research. Big data and data mining technology get attention. With the rise of social media, social media analysis has become a hot topic. Through visual analysis of social media data, libraries can understand users' perceptions of their services and start to pay attention to the application of big data and machine learning technology in library data analysis.

Boom period: 2018-2023.

In the last five years, big data has been paid more attention and mined. Data visualization is increasingly being used for predictive analytics and to build user profiles. More application research has been carried out to adapt to the goal of intelligent library service users. Through data visualization, mining of associated data, and extensive application of

bibliometric technology, libraries can predict future user behaviors and needs, and use data visualization to allocate resources, optimize services, formulate policies, and so on to further optimize services.

6.4 Development Trend of Data Visualization Application Research in Library Industry

From the initial application of data analysis technology in computer science to the library industry, gradually to the discussion of digital library, with a large number of data visualization software used in research, library researchers have applied data visualization technology in more and more service fields. The research focus of data visualization application in library industry has gradually shifted from data analysis to data service. Libraries are paying more attention than ever to the cultivation of user data literacy, the digitization of knowledge services, the

visualization of information at a glance, digital humanities research that brings digital theory and practice to the humanities, and the opportunities for innovative services and personalized services that data mining technologies may generate^[15]. Data visualization technology opens a window for users to gain insight into library knowledge resources with data, and speeds up the process of transformation from digital library to smart library.

7. Conclusion

The application of data visualization technology is the need for the transformation and development of smart library. Through data visualization, the content expressed by the data can be easily understood while ensuring the effective transmission of information. Through data analysis, gradually improve and optimize the library service environment; further enhance the user experience; Policies are implemented periodically based on information provided as the data changes. Data-driven decision making is becoming increasingly important. In the process of the library from "digital" to "intelligent", more and more attention is paid to the use of real-time data visualization and other intelligent means to understand the impact of subtle changes on the overall situation, so as to comprehensively integrate the information resources inside and outside the library.

References

- [1] Ren Ni, Wu Qiong, Li Huiquan. Analysis and Research of data visualization Technology. *Electronic Technology and Software Engineering*, 2022, No. 234 (16): 180-183. (In Chinese)
- [2] Chen Ting, Xu Shengjie, Chen Long et al. Research and Design of Chinese Patent Data Visualization Analysis System. *Computer Applications and Software*, 2023, 40 (01): 24-29. (In Chinese)
- [3] Li Wenqi, Wang Fengxiang, Sun Xianbin et al. Disordered directory data integration and visualization of. *Journal of Chinese library*, 2023, 49 (01): 82-98. The DOI: 10.13530/j.carol carroll nki jlis. 2023006.
- [4] Bakhshayesh, S.M., Ahmadi, A. and Mohebi, A. (2021), "A new subject-based retrieval and search result visualization approach for scientific digital libraries", *The Electronic Library*, Vol. 39 No. 4, pp. 572-595. <https://doi.org/10.1108/EL-08-2020-0243>.
- [5] Huang Hao. A brief introduction to the design and implementation of medical data visualization on large screen using Python+Flask+ECharts. *Digital technology and applications*, 2022, 40 (9): 200-202. The DOI: 10.19695 / j.carol carroll nki cn12-1369.2022.09.60.
- [6] Hai-yan ren. Application and development trend in the field of visual data mining research. *Journal of digital technology and applications*, 2018, 4 (8): 225 + 228, DOI: 10.19695/j. carol carroll nki cn12-1369.2018.08.120.
- [7] Yu Defeng. Analysis on data visualization services of Ivy League University libraries. *Libraries*, 2022, No. 334 (07):79-86+100.
- [8] Wang Xiuli, Gao Kai, Jiang Jihong, Yu Jing. The role of reading promotion in Library construction under network environment. *China Science and Technology Journal Database Research*, 2022 (5): 0179-0182.
- [9] Liu Xun, Zhang Yanhua, Zhang Dexiang, Li Jin. A Comprehensive Risk Evaluation Method for Digital Library Construction Based on the Perspective of Project Management. *Journal of Qingdao University (Natural Science Edition)*, 2013, 26 (03): 59-63.
- [10] Datig, I. and Whiting, P. (2018), "Telling your library story: tableau public for data visualization", *Library Hi Tech News*, Vol. 35 No. 4, Pp. 6-8. <https://doi.org/10.1108/LHTN-02-2018-0008>.
- [11] Fu Tianxin, ZHANG Qinghua, Liu Chengbin et al. Portraits of big data horizon, the university library users to construct and express visual. *Journal of Beijing agricultural college*, 2023 ((01): 90-96. The DOI: 10.19444 / j.carol carroll nki. 1671-7252.2023.01.011.
- [12] HE Bin, QIN Fuqiang, FENG Juan. How to improve the readership of weibo public number of science and technology journals--Taking *Journal of Qingdao University (Natural Science Edition)* as an example. *Science and Technology Communication*, 2021, 13 (16).
- [13] Ad Hoc Column of *Journal of Qingdao University (Natural Science*

- Edition)--Construction of Complete Domestic Demand System and High-Quality Economic Development. *Journal of Qingdao University (Natural Science Edition)*, 2021, 34 (02).
- [14]Gao Kai, Yang Zongling. Research on the Innovation of College Library Management and Service Mode in the Information Age-Recommended Readings of College Library Management and Innovative Practices. *Intelligence Theory and Practice*, 2023, 46 (06): 206.
- [15]Qin Fuqiang, Shi Qiaoling. Division of Trade and Cooperation between China and the Five Central Asian Countries in the Background of "Belt and Road" — Based on trade-related index analysis. *Journal of Qingdao University (Natural Science Edition)*, 2016, 29 (02): 128-132.