Research on Visualization of international research Cooperation Results in China Library and Information Science based on Big Data

Duanwu Yan[1a*], Chenhe Yang [2b], Qiaoyu Chen [3c]

[a*] Yanwu123@sina.com, [b] 1198435983@qq.com, [c] 2276799447@qq.com

[123] School of Economics and Management, Nanjing University of Science and Technology, Nanjing 210094, China

Abstract. With the deepening of global informatization, transnational scientific research cooperation mode has played a significant role in promoting the development of library and information science. This paper collected the journals with top impact factors among the core journals of WOS from 2000 to 2021 as the research objects of this paper, combined with literature metrology, subject analysis and other methods, and carried out research on the achievements of transnational scientific research cooperation in the field of library and information by means of knowledge graph visualization. Based on the development status and evolution characteristics, the research hotspot and knowledge transfer rule in this field in recent years are studied and summarized, and the future development direction of the subject is predicted. It is found that the scale and scope of the cooperation between China and other countries in library and information science are constantly expanding, and the cooperation shows a trend of a hundred flowers blooming among regions. Transnational scientific research cooperation in the field of library and information is shifting and developing towards the application of new technologies such as NLP as research tools and the research of new media environment such as the Internet.

Keywords: Big data analysis; library and information science; knowledge map construction; visualization topic analysis

1   Introduction

Scientific research cooperation is a cooperative and mutual assistance scientific activity carried out by researchers in order to achieve research objectives [1]. In recent years, with the deepening of information technology, scientific research cooperation has become one of the important ways for scholars to conduct scientific research [2]. Library and Information Science (LIS), as a subject that takes data as its research object, is of great significance to the mining of its research hotspots [3]. The research on the mode of scientific research cooperation is of great significance for improving the quality of scientific research and accelerating the speed of knowledge flow [4].

In previous studies, scholars mostly combined research cooperation with industrial hot spot phenomena, studied the cooperation value and influencing factors of research cooperation, and rarely analyzed this issue from the perspective of scientific research achievements. By com-

bining knowledge graph visualization, they obtained the development status and hot spot of target disciplines and predicted the development of disciplines [5][6].

Therefore, this paper will start with LIS transnational research cooperation papers to build a cooperative network map of scientific research achievements in this field. In addition, the research content in LIS field was summarized by subject analysis method, and the relevant characteristic indexes of the map were summarized. The research hotspot of transnational scientific research cooperation in LIS field was studied and the future development law was predicted.

2   Research design and data preprocessing

Paper is one of the important output forms of scientific research cooperation. By analyzing the achievements of international scientific research cooperation in the field of library and information in China, this paper makes further research on the cooperation scale, hot spot changes and development direction of China's international scientific research cooperation.

In this paper, the core database of Web of Science is used as the data source, and the 2022 Journal Impact factor published by WOS is used as the reference. All the papers published by the journal from 2000 to 2021 and included by the core database of WOS with Chinese participation are retrieved. The total number of papers in WOS core library is 6616. In this paper, Mysql database and python are used for data preprocessing. By extracting and cleaning the field data of the catalogue, a total of 1018 literatures are exported for research. In this paper, "Peoples R China" and "Taiwan" are regarded as two different regions belonging to China for analysis. In the country/region analysis, Taiwan region in this paper refers to Taiwan of China, and mainland China refers to all regions of China except Taiwan, including Hong Kong and Macao.

3   Results and discussion

3.1   Statistical analysis of the quantitative characteristics of LIS international research cooperation countries

In order to understand the scientific research situation of countries and institutions in the transnational scientific research cooperation network in the field of library and information in China at different stages, we counted the top 5 countries/regions in the volume of international cooperation with China at four different stages and the top 10 institutions in the volume of international cooperation.

Firstly, the dominant position of LIS research cooperation between mainland China and Taiwan has changed greatly in the past 20 years. The Chinese mainland has been dominant in scientific research cooperation, and the number of collaborative results has been increasing, while the proportion of Chinese Taiwan participating in the cooperation has been declining.

Secondly, the cooperation between mainland China and the United States is relatively close. In the past 10 years, the cooperation scale between China and the United States has achieved unprecedented development. In addition, China also has relatively stable cooperative relations

with Singapore, Canada and Australia, but the amount of cooperation is far lower than that with the United States.

Finally, in the past 20 years, China's scientific research institutions have become increasingly close to international cooperation. In addition, we also observed that the proportion of scientific research cooperation institutions in mainland China is increasing year by year, which can be seen that Chinese scientific research institutions further enhance the importance of international cooperation mode. In the past 10 years, scientific research institutions in other provinces and cities in China have also gradually increased the amount of international cooperation, showing the distribution of a hundred flowers.

3.2   Map of LIS international research cooperation and its visualization

In this paper, the keywords of scientific research cooperation papers in the field of library and information in China in the past 20 years are extracted and the keyword co-occurrence map is constructed, the important fields of China's international scientific research cooperation are summarized, and the clustering themes and the keywords that are studied frequently in different stages are sorted out. The author drew the keyword co-occurrence map and carried out visualization processing. The drawing results are shown in Figure 1. Table 1 shows the 6 categories of topics in the clustering diagram and the high-frequency keywords in each topic.
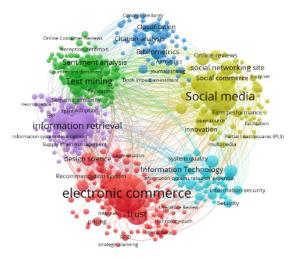

Fig. 1. Key words Co-occurrence Map of LIS international Research cooperation in China

Table 1. Graph clustering topics and high frequency keywords

Theme	High-frequency keywords	Keywords quantity
NLP and machine learning	Text mining, sentiment analysis, data mining, acceptance models, natural language processing	191
Information organization and information management	Information retrieval, Case Studies, Knowledge Sharing, Knowledge Management, Project management	160
Electronic commerce and user services	E-commerce, Trust, Design science, user habits, website design	138

Teamwork and information behavior	Social networks, social networking sites, innovation, social influence, online communities	80
Bibliometric and scientific evaluation	Bibliometrics, citation analysis, classification, academic communication, Altmetrics	64
Information system and Network Security	Information Technology, technology adoption, System quality, software development, information security	47

As can be seen from Figure 1 and Table 1, There are six main research topics in our international scientific research cooperation, among which "NLP and machine learning" is the most abundant, while "bibliometrics and scientific evaluation" and "information system and network security" are less researched fields in China's international scientific research cooperation..

The research content of NLP(natural Language processing) and machine learning (topic 1) combines relevant theories and methods of machine learning, and regards knowledge related to the field of computer as a data processing and analysis tool [7]. Information organization and Information Management (topic 2) focuses on the organization and management of information resources as well as their application [8]. The research content of E-commerce and User Service (topic 3) is centered on e-commerce, and many qualitative and quantitative research methods and models are used to improve the quality and quantity of transactions [9]. The research content of team cooperation and Information Behavior (topic 4) focuses on the communication and cooperation between individuals and the information behavior carried out to meet users' goals [10]. Bibliometrics and Scientific Evaluation (topic 5), on the other hand, used more scientometrics methods to analyze the literature by extracting relevant data such as reference relationship and title data of scientific research literature. Information system and Network Security (topic 6) mainly aims to improve the performance of information related software facilities and network environment to ensure a safer, more reliable and more efficient working environment.

First of all, four words of information organization and Information management (topic 2) have a high research frequency, which indicates that this topic is deeply concerned by the majority of scholars, and further proves that information is one of the key research objects in library and information science. In addition, we can find that "e-commerce" is in a higher position in each period, and the frequency of occurrence is relatively balanced. It can be seen that "e-commerce" has always been one of the key studies of transnational cooperation in the field of library and information in China.

Secondly, in the first 10 years, the word "information retrieval" always had a high ranking, and in the second 10 years, the frequency of this word decreased significantly, on the contrary, the word "social media", which is not difficult to see the migration and changes of the research hotspot of transnational scientific research cooperation in the past 20 years. Scholars can move their research centers to fields that focus on human subjects such as "social media" and "information behavior", and dig information behavior rules in combination with new media environments such as mobile phones and the Internet.

Finally, combined with the high-frequency words in each stage and their corresponding topics, we can observe that the proportion of information organization and information management (topic 2) basically shows a decreasing trend, while that of teamwork and information behavior (topic 4) has achieved good development in the past 10 years. NLP and machine learning (top-

ic 1) occupy a large proportion in the last five years. From this, we can see that scholars have gradually shifted their focus to the research on users and the network environment, compared with the traditional pure research object of information. With the progress of computer technology, more and more scholars choose to conduct scientific research with computer technology, and the research direction in the field of library and information has begun to move towards a more scientific and humanistic direction.

4   Conclusion

Based on the change of the number of papers in the field of library and information and the research content of the cooperative papers in the past 21 years, this paper constructs and visualizes the map, obtains the key points and hot spots of the research content of transnational research in the field of library and information in China, and predicts the future development hot spots and directions of this field.

It is found that the scale of China's national/regional and institutional cooperation network continues to expand, and the scale of national scientific research cooperation led by China continues to radiate to neighboring countries, and attracts new countries from different regions to join the cooperation network. In addition, with the development of Internet environment and computer technology, the research content of library and information science is more and more technical, and more attention is paid to the role of new media in information behavior. Based on the above conclusions, on the one hand, Chinese research institutions should further improve the transnational cooperation system, actively build cooperation platforms, and encourage scholars to cooperate and exchange. We can set up an inter-institutional base for cooperation, and also can increase funding projects for transnational research cooperation, and reward scholars who achieve academic achievements through transnational research cooperation. On the other hand, scholars should actively learn advanced subject technologies, closely connect with the Internet environment and information technology to bring changes to library and information science, and jointly promote the progress and development of the field.

There are still some limitations in this study, and the research samples need to be enriched. In this study, the author only selected periodicals with relatively high impact factors in the field of library and information in the WOS database for data collection. The author believes that a larger volume of periodicals and papers can be selected for research in the subsequent study, which will enrich the experimental data and make the experimental results more complete.

References

[1] Luukkonen, T., Tijssen, R., Persson, O., Sivertsen, G. (1993). The measurement of international scientific collaboration. scientometrics, 28(1): 15-36. https://doi.org/10.1007/bf02016282.

[2] Wang, P., Wang, N.Z. (2021). An Analysis of the hot topics of publishing Research and the Development Trend of Publishing Industry in China from 1998 to 2020: Based on the econometric study of high-impact papers in source journals of CSSCI. View on Publishing, 400(22): 49-52+88. DOI:10.16491/j.cnki.cn45-1216/g2.2021.22.011.

[3] Zhang, X.Y., Song, W., Song X.Y. (2015). Evolution Analysis of International Cooperation Network of Chinese scientific research Institutions: Taking SCI innovation management papers as an example. China University Science & Technology, 323(07): 12-17. DOI:10.16209/j.cnki.cust.2015.07.003.

[4] Chen, J., Zhang, F. (2022). Research on research cooperation Situation of Chinese public libraries based on bibliometric analysis: A case study of 48 sub-provincial public libraries and above. Library Work and Study, 315(05):35-42+75. DOI:10.16384/j.cnki.lwas.2022.05.012.

[5] Larivière, V., Gingras, Y., Sugimoto, C.R., Tsou, A. (2015). Team size matters: Collaboration and scientific impact since 1900. Journal of the Association for Information Science and Technology, 66(7): 1323-1332. https://doi.org/10.1002/asi.23266.

[6] Wang, NN., Xiong, H.X., Liu, M.H. (2022) Recommendation of Scientific Research Collaborators Based on Multidimensional Decision Attributes. Information Science, 40(07): 93-101. DOI:10.13833/j.issn.1007-7634.2022.07.012.

[7] Laband, D.N., Tollison, R.D. (2000). Intellectual collaboration. Journal of Political economy, 108(3): 632-662. https://doi.org/10.1086/262132.

[8] Larivière, V., Gingras, Y., Sugimoto, C.R., Tsou, A. (2015). Team size matters: Collaboration and scientific impact since 1900. Journal of the Association for Information Science and Technology, 66(7): 1323-1332. https://doi.org/10.1002/asi.23266.

[9] Liu, Y.M., Zang, S., Wang, D.D. (2022) Superiority Research Direction for First-Level Disciplinary Institution in Library Information and Archives Management. Information Science, 40(03):174-182. DOI:10.13833/j.issn.1007-7634.2022.03.022.

[10] Yan, W.W., Wen, X., Cao, C.Y. (2021) Thematic Analysis of High Web Impact LIS Research Publications in Academic Social Networks——Taking ResearchGate as an Example. Library Tribune, 41(03):50-57. https://kns.cnki.net/kcms/detail/44.1306.G2.20201015.2037.004.html