Topic Popularity Acceleration Index: A Novel Method for Identifying Disciplinary Research Hotspots (Postprint)

Authors: Rong Guoyang, Li Changling, Fan Qingqing, Guo Fengjiao

Date: 2023-04-01T00:00:00+00:00

Abstract

[Purpose/Significance] Synthetically identifying research hotspots across different temporal stages from both horizontal and vertical perspectives facilitates grasping the developmental trajectory and direction of a discipline, thereby expanding research avenues for hotspot topic identification methodologies. [Method/Process] We construct a cumulative topic popularity model TP to reflect the horizontal relative research popularity of topics within the discipline; a topic popularity acceleration index model TAI to quantify vertical velocity changes in research topic development; and a disciplinary research hotspot identification model TP*TAI that comprehensively reflects topic research popularity and its variations from both horizontal and vertical dimensions. [Results/Conclusions] Employing research literature from the library and information science domain spanning 2001-2020 as a sample for empirical analysis, the results demonstrate that: the model can effectively identify research hotspots at various temporal stages, classify them into three categories—emerging, stable, and declining—thereby achieving a dynamic depiction of disciplinary research.

Full Text

Topic Acceleration Index: A New Method for Identifying Discipline Research Hotspots

Rong Guoyang, Li Changling, Fan Qingqing, Guo Fengjiao
Institute of Information Management, Shandong University of Technology, Zibo 255049

Abstract

[**Purpose/Significance**] Comprehensively identifying research hotspots from both horizontal and vertical perspectives across different time stages helps grasp the development trajectory and direction of a discipline, expanding research ideas for hotspot identification methods. [**Method/Process**] This study constructs the Cumulative Topic Popularity model (TP) to reflect the horizontal relative research heat of a topic within a discipline; builds the Topic Acceleration Index model (TAI) to quantify the longitudinal speed changes in topic development; and establishes the discipline research hotspot identification model (TP*TAI) to comprehensively reflect topic research heat and its changes from both horizontal and vertical dimensions. [**Result/Conclusion**] Using research literature in library and information science from 2001–2020 as a sample for empirical analysis, the results demonstrate that the model can effectively identify research hotspots at various time stages and classify them into three types: frontier, stable, and declining, thereby achieving dynamic description of disciplinary research.

Keywords: topic popularity; research hotspots; identification method; TP*TAI model

1 Introduction

A paper' s keywords or subject terms condense and distill its core content, representing the research topic. If a keyword or subject term appears repeatedly in literature within its discipline, it indicates that the topic represented by that keyword is a research hotspot in the discipline and a high-heat research topic widely concerned and studied by researchers. Different time stages in a discipline have different research hotspots, which represent the focus and development direction of disciplinary research at those stages. Identifying disciplinary hotspots and their changes helps researchers grasp the development trajectory and direction of the discipline, select research topics, and allocate research resources rationally, while also providing a research foundation for hotspot topic evaluation, future trend prediction, and related law discovery. Currently, constructing scientific methods to identify disciplinary research hotspots has become key to grasping the current state of a discipline and predicting future trends. Many scholars at home and abroad have chosen different methods from various perspectives to identify disciplinary research hotspots.

First, keyword burst, co-occurrence, and citation analysis methods are used to identify disciplinary research hotspots. K. Mane and K. Börner employed Kleinberg burst detection algorithms and co-word analysis to identify hot topics such as gene mutation, molecular sequences, and protein research in PNAS papers. Liu Xiaohui et al. improved the TF-IDF algorithm to identify research hotspots in information science in 2015, including user research and big data. Gao Jiping et al., Tang Qiang et al., and Hu Xiumei et al. used multi-word co-occurrence and cluster analysis to identify research hotspots in different disciplinary fields such as digital information transmission, 3D printing technology, and library and information science strategic planning. Y. Chang et al., X. Ping, D. Rossetto et al., and C. Jebbari et al. used citation coupling and co-citation analysis to identify research hotspots in library and information science, international anticancer research, business management, biomedicine, and other fields.

Second, time series analysis or time-weighted methods are used to identify disciplinary research hotspots and their change trends. Xiao Tingting et al. employed time series visualization analysis and found that semantic annotation research is based on ontologies and knowledge bases, developing continuously with in-depth semantic web research. Liu Ziqiang et al. proposed a research hotspot evaluation and prediction method based on time series models, analyzing the status and development trends of competitive intelligence research from 2005–2014, predicting 2015 research targets and verifying them, finding the model feasible and effective. Zhou Xin et al. constructed a keyword frequency change rate model to analyze CSSCI-indexed information science journals from 2000–2014, identifying growth, stable, and declining research hotspots. Li Changling et al. improved the z-index based on time factors, comparing ranking changes in topic heat before and after the improvement, dividing information science research hotspots from 2014–2018 into rising, stable, and declining types. Feng Guohe et al. constructed a time-weighted keyword frequency analysis model to identify research hotspots in CSSCI library and information science journals and classified results into rising, declining, and stable keywords. J. Li et al. proposed clustering coefficient and path length calculation methods based on time-perspective co-word networks to identify research hotspots in library and information science from 2014–2016, including big data, Altmetrics, and mobile library categories.

Finally, computer algorithms and models are used to identify disciplinary research hotspots. Sun Haisheng improved co-word and co-citation analysis methods based on hypernetwork models to identify library and information science research hotspots from 2014–2016, including 7 categories such as big data, Altmetrics, and mobile libraries. C. Figuerola et al. and X. Han applied topic modeling statistical techniques and latent Dirichlet allocation models to identify hot topics in different disciplines. Ruan Guangce et al. used the Doc2Vec method to vectorize paper abstracts, analyzed their similarity, generated hotspot topic paper collections, extracted topic descriptions, and identified education hotspots including higher education reform, higher education equity, and learning methods. Qiu Huilin et al. integrated LDA2vec models with Word2Vec and document vectorization models to propose a multi-source data hotspot identification method, identifying machine learning hotspots including text classification and feature detection.

In summary, current research hotspot identification mainly employs methods such as word frequency burst, co-occurrence clustering, citation analysis, word frequency change rate, time weighting, and model algorithms to identify disciplinary hot research topics from certain perspectives, but does not deeply analyze the changes in topic heat. Therefore, this paper analyzes the longitudinal changes in heat acceleration while analyzing the horizontal heat of disciplinary research topics to comprehensively identify research hotspots at different stages. The steps are as follows: Construct the cumulative topic popularity model TP, i.e., the proportion of cumulative word frequency of a research topic within the discipline, reflecting the horizontal relative research heat of the topic within a certain time period; Construct the topic acceleration index model TAI, representing the change in topic heat growth rate, reflecting the longitudinal heat and its changes of the research topic from a temporal perspective; Construct the discipline research hotspot identification model TP*TAI, comprehensively reflecting the horizontal relative research heat and longitudinal changes of research topics within the discipline to identify research hotspots in each time period; Establish classification criteria to subdivide hotspot topics into three types: frontier, stable, and declining, to accurately grasp disciplinary development dynamics.

---

2 Discipline Hotspot Identification Based on Topic Heat Acceleration

2.1 Cumulative Topic Popularity Model (TP) The Cumulative Topic Popularity (TP) model uses the proportion of cumulative word frequency (i.e., research literature volume) of a research topic in a certain time period within the total disciplinary literature volume, expressed as:

$$TP = \frac{\sum_{t=n}^{m} C_t}{\sum_{t=n}^{m} P_t} \quad (n \le i \le m) \tag{1}$$

where t represents the year, C_t is the research literature volume of a certain topic in year t, and P_t is the total disciplinary literature volume in year t. n is the first year the topic appears or the starting year of data grouping, and m is the most recent year or the cutoff year of data grouping.

In formula (1), if $m = n$, it represents the relative heat of a topic in year t. The TP model uses relative cumulative volume to measure research topic heat year by year, which can not only reflect the heat of a topic' s development up to year t but also eliminate errors caused by different absolute literature quantities in each year. Although this model can reflect the horizontal relative research heat of a topic in a certain time period within the discipline, the cumulative calculation method cannot express the changing trend of topic heat. Therefore, this paper constructs the Topic Acceleration Index model (TAI) to discover the longitudinal change trend of topic research heat.

2.2 Topic Acceleration Index Model (TAI) In physics concepts, acceleration refers to the ratio of velocity change to the time taken for this change, expressed as:

$$a = \frac{v_2 - v_1}{\Delta t} \tag{2}$$

where $v_2 - v_1$ refers to the velocity change, and Δt refers to the time interval between two velocity changes. If acceleration a is greater than zero, the object is accelerating in the positive direction; if acceleration a is less than zero, the object is decelerating in the positive direction; if acceleration a equals zero, the object is stationary or moving at constant velocity.

This paper introduces the acceleration concept into research hotspot identification to measure the accelerated changes in topic research heat. TP_t can be regarded as the relative growth rate in year t within the cumulative heat. Then, the topic heat acceleration a with a one-year interval ($\Delta t = 1$) can be expressed as:

$$a = TP_t - TP_{t-1} \tag{4}$$

Similar to the meaning of acceleration in physics, if $a > 0$, it indicates that the topic research heat shows an accelerated growth trend; if $a = 0$, it indicates that the topic heat growth rate remains unchanged; if $a < 0$, it indicates that the topic heat growth is slowing down. Obviously, regardless of the value of topic heat acceleration a, the topic research heat is always non-negative. To express the relationship between topic heat acceleration and topic heat, this paper takes the exponential of topic heat acceleration a with base e to obtain the Topic Acceleration Index (TAI) model, expressed as:

$$TAI = e^a = e^{TP_t - TP_{t-1}} \tag{5}$$

In this formula, the TAI index is always greater than zero, consistent with the characteristic that topic research heat is always non-negative. If $a > 0$, then $TAI > 1$; if $a < 0$, then $0 < TAI < 1$; if $a = 0$, then $TAI = 1$.

Since the TAI model measures the acceleration of topic heat rather than velocity, on the one hand, it can identify research topics with low cumulative literature volume but extremely high acceleration—such topics' research heat grows sharply in a short time and may be new knowledge growth points. On the other hand, it can also identify research topics with high cumulative literature volume but negative acceleration—such topics' research heat declines rapidly in a short time and should be screened out from the hotspot list. Therefore, the advantage of this model lies in its ability to perceive short-term heat changes in research topics and effectively grasp disciplinary development trends.

2.3 Discipline Hotspot Identification Model (TP*TAI) Cumulative topic popularity TP considers the horizontal relative research heat of a topic compared with other research topics in the discipline, while the topic acceleration index TAI reflects the longitudinal heat change of a topic during disciplinary development. This paper constructs the discipline research hotspot identification model as:

$$TP * TAI = \frac{\sum_{t=n}^{m} C_t}{\sum_{t=n}^{m} P_t} \times e^{TP_t - TP_{t-1}} \quad (n \leq i \leq m) \tag{6}$$

In this model, cumulative topic popularity TP reflects the horizontal relative heat of a research topic in the discipline field during a certain time period, while the topic acceleration index TAI reflects the change in topic heat growth rate. TP*TAI combines the two to comprehensively evaluate topic research heat from both horizontal and vertical aspects. Therefore, the TP*TAI model can reflect the cumulative research heat and dynamic changes of topics, identifying stable core hotspot themes with long time spans at the macro level, and identifying stage-specific emerging hotspot themes and current frontier hotspots at the micro level.

---

3 Empirical Analysis: Library and Information Science as an Example

3.1 Data Source and Preprocessing This study takes library and information science as an example, using CNKI journal database as the data source. Sample source journals were selected based on the top 10 CSSCI journals ranked by comprehensive impact factor in 2019. Since data for *Journal of the China Society for Scientific and Technical Information* from 2003–2012 is missing in CNKI and cannot be used as a sample, the final selected sample journals include: *Journal of Library Science in China*, *Documentation, Information & Knowledge*, *Journal of Academic Libraries*, *Library and Information*, *Information Studies: Theory & Application*, *Library and Information Service*, *Information and Documentation Services*, *Information Science*, and *Journal of Intelligence*. A total of 55,444 bibliographic records were collected from these nine journals for 2000–2020. The data from 2000 was used as the first group of TP_{t-1} for TAI index calculation, and data from 2001–2020 was used for discipline research hotspot identification. Data collection time was January 12, 2021. After removing 2,127 non-research papers such as prefaces, editorial announcements, and calls for papers from the 55,444 records, 53,317 valid samples were obtained.

This paper selects author-identified keywords as the paper' s research topics to test model effectiveness and identify case discipline research hotspots. The sample literature data was preprocessed as follows: Use bibexcel to split keyword fields; Count annual keywords and their frequencies; Screen valid words. Remove words with unclear research purposes such as "influencing factors" and "analysis"; remove words indicating research background such as "United States" and "China"; merge synonymous and near-synonymous keywords, such as merging "COVID-19" and "novel coronavirus pneumonia" into "COVID-19".

3.2 Discipline Hotspot Identification The preprocessed keyword sample data was processed as follows:

(1) **Construct year-keyword frequency matrix.** A total of 1,359 keywords with annual frequency thresholds greater than or equal to 5 were selected. The corresponding data of keyword frequency and year C_t, as well as the total annual publication volume P_t, were summarized in Table 1 (partial results).

(2) **Calculate cumulative topic popularity TP.** To describe the development details of China' s library and information science discipline while avoiding calculation result distortion, the 2001–2020 data was divided into 5-year groups, totaling 4 groups. The data in Table 1 were substituted into formula (1), with n = group start time and m = group end time, to calculate TP values. Partial calculation results are shown in Table 2.

(3) **Calculate topic acceleration index TAI.** Since the relative annual growth rate TP_t of topic heat has small numerical values, $TP_t - TP_{t-1}$ lacks good discriminability, causing the topic acceleration index TAI to be close to 1. To better distinguish the impact of topic research heat changes and achieve the purpose of reflecting topic growth rate changes, TP_t was magnified by a thousand times. For convenience of calculation, formula (5) was adjusted to $TAI = e^{TP_t‰ - TP_{t-1}‰}$ to achieve the magnification of TP_t. The data in Table 1 were substituted into this formula, and partial TAI calculation results are shown in Table 3.

(4) **Discipline hotspot identification and group ranking.** The data in Table 2 and Table 3 were multiplied correspondingly to obtain 20 TP*TAI values for each topic word during 2001–2020. These 20 values were divided into 4 groups of 5 years each. The average value within each group was taken as the comprehensive score of a topic in that time period. All topic words were ranked within each group according to their comprehensive scores. Partial calculation results are shown in Table 4.

3.3 Classification of Discipline Research Hotspots Based on the ranking changes of each research topic, this paper classifies research hotspots into three categories: frontier, stable, and declining. Luo Rui et al., through concept analysis and characteristic research, believe that research frontiers have two features: recent emergence and high innovation value. Zheng Yanning et al. consider frontiers as representing the latest research progress or trends in a research field relative to specific research areas and time periods. Yan Duanwu et al. define research frontiers from the perspective of topic evolution as newly emerged and potentially developing research topics. Q. Wang believes emerging frontier topics have five characteristics: novelty, rapid growth, coherence, high impact, and uncertainty.

Based on the above views, this paper argues that frontier research hotspots should have three characteristics: recent emergence or prominence, rapid growth, and high impact. Therefore, the definitions of the three types of research hotspots are as follows: **Frontier research hotspots.** Topics ranked in the top 2% of total sample research topics in the most recent 5-year period (2016–2020) and that are prominent or appear for the first time in the most recent two groups of data, expressed as ranking rising by more than 50% in 2016–2020 or 2010–2015 compared with the previous group. **Stable research hotspots.** Topics with consistently high comprehensive scores in disciplinary research, expressed as being in the top 10% of total sample research topics in all 4 groups, with ranking fluctuations between adjacent groups not exceeding 5% of total sample research topics. **Declining research hotspots.** Topics that were hotspots in a certain time period but subsequently declined in heat, i.e., topics with comprehensive scores showing a decreasing trend across 4 groups, with a range greater than 20% of total sample research topics.

A VBA program was written to classify the discipline hotspot identification results in Table 4 according to the classification principles shown in Figure 1 [Figure 1: see original paper], dividing them into three categories. The top 8 results for each category are shown in Table 5 .

"Frontier" research hotspots are high-heat research topics in library and information science in recent years, characterized by high disciplinary research heat proportion and fast growth rate, with great research potential and momentum. "Stable" research hotspots have relatively stable heat at the macro level and fluctuating development at the micro level, representing relatively solid core research content in information science. "Declining" research hotspots show a year-by-year decreasing trend, indicating that such topics have accumulated certain research results and become relatively mature; in recent years, such research has either become more in-depth and detailed or shifted to related fields.

---

4 Model Effectiveness Analysis

4.1 Model Validity Analysis of Tables 1–4 reveals that the TP *TAI model can effectively identify discipline hotspots from both horizontal and vertical aspects. TP measures the relative heat of research topics in a certain time period, while TAI reflects the changes in topic research heat over time. Comparison with related previous studies in similar time periods shows that the TP*TAI model can not only analyze macro-level discipline hotspots over long time spans through annual cumulative data but also identify research hotspots in different time periods, including frontier, stable, and declining hotspot topics.

(1) **The cumulative topic popularity model TP can effectively reflect**

the relative heat of a topic at a certain time. For example, in rows 1 and 2 of Table 1, the absolute word frequency of "library" in 2017 and "information service" in 2008 and 2009 is the same (74). However, due to different total disciplinary research literature volumes, the TP values reflecting cumulative topic heat differ significantly. In rows 1 and 2 of Table 2, the cumulative topic popularity TP of "library" in 2017 is 0.038, while that of "information service" in 2008 and 2009 is 0.027. Thus, this model eliminates measurement errors caused by different total literature volumes each year and can reflect the cumulative heat of a topic up to the measurement year.

(2) **The topic acceleration index model TAI can effectively reflect the changes in a topic' s research heat over time.** In row 6 of Table 3, the TAI values of "knowledge management" in 2002, 2004, and 2005 are far greater than 1, reaching the ten-thousand level, representing extremely large acceleration in topic heat during these three years, indicating a surge in research heat for "knowledge management."However, starting from 2008, except for 2015 and 2017, the TAI values of this research topic are all less than 1, representing negative acceleration in research heat, indicating that the growth rate of "knowledge management" declined during this period and entered a declining phase. Thus, the acceleration dimension introduced by this model can effectively identify the growth rate of topic heat in different time periods, compensating for the TP model' s inability to reflect longitudinal development changes in topic heat, effectively demonstrating accelerated changes in topic research heat and achieving dynamic analysis.

(3) **The TP*TAI model integrates the two indicators of cumulative topic popularity TP and heat acceleration TAI, measuring topic heat and its changes from both horizontal and vertical perspectives to effectively identify research hotspots in each period.** For example, in Table 2, the TP value of "big data" in 2011–2015 is significantly smaller than that of "digital library," indicating that the research heat of "big data" in information science during this period was lower than that of "digital library." In Table 3, the heat acceleration TAI of "big data" in 2012–2015 is far greater than that of "digital library," indicating that the research momentum and heat trend of "big data" were higher than those of "digital library." Combining horizontal and vertical performance, in Table 4, the TP*TAI comprehensive ranking of "big data" in 2011–2015 is higher than that of "digital library," ranking 1st and 4th respectively. Additionally, as a high-heat research topic that first appeared in 2020, "COVID-19" has large TP and TAI values in Tables 2 and 3, representing both high research heat and high growth rate. The TP*TAI calculation result in Table 4 is as high as 6.81×10^{16}, making it the hottest research topic of that year. Therefore, the two indicators in the TP*TAI model can effectively influence the comprehensive ranking, and top-ranked topic words in each time stage have two characteristics: high proportion of topic research literature volume in total disciplinary

literature volume (i.e., high cumulative topic popularity) and high growth trend (i.e., large topic heat acceleration).

4.2 Model Advantages This paper divides the 20-year span of library and information science research literature (2001-2020) into four 5-year periods to comprehensively identify discipline hotspots from both macro and micro perspectives. To further analyze model effectiveness and discover its advantages, research literature from similar time periods was selected for comparative analysis. Since no relevant literature exists on library and information science research hotspots for 2016-2020, three similar periods (2001-2005, 2006-2010, and 2011-2015) were selected for comparative analysis with similar studies. Relevant data are shown in Table 6 .

The research conclusions on different stages obtained by previous scholars through word frequency or co-occurrence analysis are the same as or similar to the results of this paper in the same time periods (Table 4), indicating that the TP*TAI identification model in this paper is feasible and effective. At the same time, compared with previous studies, this paper refines research granularity by time period, comprehensively reflecting topic research heat and trends from both horizontal and vertical aspects, with the following advantages:

(1) **It can analyze micro, meso, and macro-level disciplinary development changes.** By subdividing the time span, this paper can grasp the distribution of disciplinary research hotspots in each time period on the one hand, and understand the development trajectory and course of the discipline over long time spans on the other. For example, Tables 2 and 3 can analyze annual disciplinary hotspots and their changes; Table 4 can identify research hotspots in different time periods and grasp the stage-specific development and changes within the long sample time span.

(2) **It can identify emerging research themes with low disciplinary proportion but fast growth rate in each time period, and can also identify and screen out gradually declining research themes with high disciplinary proportion but slow growth rate.** In 2001-2005, compared with contemporary literature [28], this paper not only covers the identification results of previous research but also identifies emerging hotspot topics such as WTO/intellectual property/e-commerce. These prominent topics were influenced by the research environment: the "Capital E-commerce Project" was launched in 1998, the China E-commerce Association was established in Beijing in 2000, and China joined the WTO in 2001. Although the total research literature volume had a low disciplinary proportion at that time, the research heat growth rate was fast, so it could be well identified by this paper.

In 2006-2010, some identification results of this paper are the same as contemporary literature [29], such as web2.0/information resources/library and information science research. This paper also supplements and identifies different research hotspots in the same period, such as: the library and information field began research on "ontology" in 2004, and the research heat increased rapidly during 2006–2010, so this paper effectively identified "ontology" as an emerging topic. Meanwhile, subject librarians/virtual reference consultation are traditional library service research fields with high literature volume proportion but low research heat growth rate, so they were identified and screened out by this paper.

In 2011–2015, compared with contemporary literature [30], in addition to covering previous research results, this paper also identified hotspot topics such as big data. In March 2012, the Obama administration released the "Big Data Research and Development Initiative." In 2015, Wuhan University held the international symposium on "Big Data Era Library and Information Science Theory and Education Development Strategies." The library and information discipline, with information data management as its core research content, quickly introduced big data research, and its heat surged, making "big data" rank first in research heat during this period.

(3) **It can identify current frontier hotspot topics in the discipline.** The COVID-19 pandemic swept the globe in 2020. In 2017, the National Social Science Fund major project "Research on the Construction of Information Science Discipline and Future Development Path of Information Work" was launched. In 2018, the General Office of the State Council issued and implemented the "Scientific Data Management Measures." Therefore, this paper identifies COVID-19/public health emergencies/information work/scientific data as hotspot topics in 2016–2020 and frontier issues in library and information science. The sustained growth of research heat in big data/artificial intelligence remains a frontier hotspot in library and information science.

In summary, through literature data verification and expert consultation, the research hotspots identified in each period in this paper are relatively consistent with the actual development of library and information science, indicating that the model is feasible and effective.

This paper constructs the discipline research hotspot identification model TP*TAI, where the cumulative topic popularity model TP reflects horizontal relative hotspots within the discipline, and the topic acceleration index model TAI reflects longitudinal heat changes from a temporal perspective. Empirical evidence shows that the model has the following characteristics: It reflects the relative heat and changes of research topics at each time stage from both horizontal and vertical perspectives, helping to grasp the medium- and long-term development course and direction of the discipline; It effectively identifies "frontier," "stable," and "declining" research hotspots in library and information science over the past 20 years, achieving dynamic analysis of disciplinary research.

Although the model has been effectively validated in the field of library and information science, future research should: on the one hand, verify and improve the model from multiple dimensions to enhance its applicability to other disciplines; on the other hand, extract keywords representing paper research topics from titles, abstracts, and full texts to make discipline hotspot identification more comprehensive and effective.

References

[1] Qiu Junping, Wen Fangfang. Visual analysis of research hotspots and frontiers in library and information science in the past five years: a bibliometric study based on 13 high-impact foreign source journals[J]. Journal of Library Science in China, 2011, 37(2): 51-60.

[2] Luo Rui, Xu Haiyun, Dong Kun. Review of domain frontier identification methods[J]. Library and Information Service, 2018, 62(23): 119-131.

[3] GLÄNZEL W, THIJS B. Using 'core documents' for detecting and labelling emerging topics[J]. Scientometrics, 2012, 91(2): 399-416.

[4] TU Y, SENG J. Indices of novelty for emerging topic detection[J]. Information Processing & Management, 2012, 48(2): 303-319.

[5] MANE K, BÖRNER K. Mapping topics and topic bursts in PNAS[J]. Proc Natl Acad Sci USA, 2004, 101(S1): 5287-5290.

[6] Liu Xiaohui, Li Changling, Feng Zhigang. Analyzing discipline research hotspots based on improved TF-IDF method: taking information science as an example[J]. Information Science, 2017, 35(7): 82-85, 98.

[7] Gao Jiping, Ding Kun, Pan Yuntao, et al. Implementation of multi-word co-occurrence analysis method and its application in research hotspot identification[J]. Library and Information Service, 2014, 58(24): 80-85.

[8] Tang Qiang, Wang Yamin, Zhao Yan. Research hotspots and collaborative groups based on g-index and co-occurrence index[J]. Journal of Intelligence, 2014, 33(9): 72-75.

[9] Hu Xiumei, Gao Fan. Research hotspots of library strategic planning in domestic library and information field[J]. Library and Information Service, 2016, 60(9): 13-17, 27.

[10] CHANG Y, HUANG M, LIN C. Evolution of research subjects in library and information science based on keyword, bibliographical coupling, and document co-citation visualization analysis[J]. Scientometrics, 2015, 105(1): 611-622.

[11] XIE P. Study of international anticancer research trends via co-citation and co-citation analyses[J]. Scientometrics, 2015, 105(3): 2071-2087.

[12] ROSSETTO D, BERNARDES R, BORINI F, et al. Structure and evolution of innovation research in the last 60 years: review and future research agenda

based on bibliometric and co-citation visualization analysis[J]. Scientometrics, 2018(115): 1329-1363.

[13] JEBBARI C, VIEDMA E, COBO M. The use of citation context to detect the evolution of research topics: a large-scale analysis[J]. Scientometrics, 2021, 126(4): 2971-2989.

[14] Xiao Tingting, Qiu Junping, Zu Xuan, et al. Knowledge graph analysis of research hotspots and evolution of semantic annotation[J]. Information Studies: Theory & Application, 2015, 38(1): 1-6, 22.

[15] Liu Ziqiang, Wang Xiaoyue, Bai Rujiang. Research on research hotspot analysis and prediction method based on time series model[J]. Information Studies: Theory & Application, 2016, 39(5): 27-33.

[16] Zhou Xin, Chen Yuanyuan. Analysis of discipline research development trends from the perspective of keyword frequency change: taking domestic information science research as an example[J]. Journal of Intelligence, 2016, 35(5): 133-140.

[17] Li Changling, Pai Yanxin, Xiang Fuzhong, et al. Discussion on identification method of highly cited discipline research hotspots based on improved z-index[J]. Information Studies: Theory & Application, 2020, 43(6): 69-75, 96.

[18] Feng Guohe, Kong Yongxin. Discipline hotspot research based on time-weighted keyword frequency analysis[J]. Journal of the China Society for Scientific and Technical Information, 2020, 39(1): 100-110.

[19] LI J, WANG Y, YAN B. The hotspots of lifecycle assessment for bioenergy: a review by social network analysis[J]. Science of the total environment, 2018, 625: 1301-1308.

[20] Sun Haisheng. Research hotspot detection and clustering topic description based on hypernetwork model[J]. Journal of Intelligence, 2017, 36(6): 93-98.

[21] FIGUEROLA C G, MARCO F J G, PINTO M. Mapping the evolution of library and information science (1978-2014) using topic modeling on LISA[J]. Scientometrics, 2017, 112: 1507-1535.

[22] HAN X. Evolution of research topics in LIS between 1996 and 2019: an analysis based on latent Dirichlet allocation topic model[J]. Scientometrics, 2020(125): 2561-2595.

[23] Ruan Guangce, Xia Lei. Journal paper hotspot topic identification based on Doc2Vec[J]. Information Studies: Theory & Application, 2019, 42(4): 107-111.

[24] Qiu Huilin, Shao Bo. Research on scientific research hotspot identification method in multi-source data environment[J]. Library and Information Service, 2020, 64(5): 78-88.

[25] Zheng Yanning, Xu Xiaoyang, Liu Zhihui. Research on research frontier

identification method based on keyword co-occurrence[J]. Library and Information Service, 2016, 60(4): 85-92.

[26] Yan Duanwu, Su Qiong, Zhang Xinyue. Research on scientific domain frontier detection based on temporal topic association evolution[J]. Information Studies: Theory & Application, 2019, 42(7): 144-150.

[27] WANG Q. A bibliometric model for identifying emerging research topics[J]. Journal of the association for information science and technology, 2018, 69(2): 1-25.

[28] Qiu Junping, Zhou Chunlei, Yang Siluo. Review and prospect of information science research in China during the 30 years of reform and opening up (III): analysis of development stages and trends of information science[J]. Library and Information Studies, 2009, 2(3): 1-9.

[29] Wang Lanjing. Research hotspots and key fields of library, information and archive management discipline in China from 2004-2009: analysis based on keywords of CSSCI source literature[J]. Library and Information Service, 2011, 55(16): 68-71.

[30] Wang Zhijin, Li Boya. Analysis of research hotspots and problems in Chinese information science: based on core information science journals from 2010-2014[J]. Information Studies: Theory & Application, 2016, 39(9): 7-13.

[31] Qiu Junping, Ma Xiujuan. Bibliometric analysis of domestic e-commerce research papers from 1998-2009[J]. Information Science, 2011, 29(5): 641-646.

[32] Li Jiankang, Zhang Chunhui. Progress in ontology research and its applications[J]. Library Tribune, 2004(6): 80-86.

[33] Yang Yi. Bibliometric study of "big data" [J]. Library Journal, 2012, 31(9): 29-32.

[34] Guo Xiaowan, Ran Congjing, Wu Dan, et al. Library and information science theory and education development strategies in the big data era: summary of the 4th China-US International Symposium on Library and Information Science Education in the Digital Age[J]. Library and Information Knowledge, 2016(1): 116-121.

[35] General Office of the State Council. Notice of the General Office of the State Council on issuing the Measures for the Management of Scientific Data[EB/OL]. [2021-05-23]. http://www.gov.cn/zhengce/content/2018-04/02/content_{5279272}.htm.

Author Contributions

Rong Guoyang: Research method design, data processing, paper writing and revision;

Li Changling: Paper framework design, paper guidance, paper writing and revision;
Fan Qingqing: Data collection, paper revision and proofreading;
Guo Fengjiao: Paper guidance, paper revision and proofreading.

Publication Ethics Statement

The following statement appears to be a journal-specific publication ethics declaration and is included here for completeness:

To strengthen and enhance academic norms, research integrity, and academic ethics in the processes of academic paper writing, review, and editing, establish good academic atmosphere, promote scientific spirit, resolutely resist academic misconduct, and establish and maintain a fair, just, and open academic exchange ecological environment, the Magazine House of *Library and Information Service* (including the editorial departments of *Library and Information Service* and *Knowledge Management Forum*) has formulated a publication ethics statement based on the actual conditions of the two journals, which was officially released in February 2020.

This publication ethics statement commits that the two journals will strictly abide by and implement national policies and regulations related to academic ethics and editing and publishing, standardize the behavior of authors, peer review experts, and journal editors throughout the entire editing and publishing process, and accept supervision from the academic community and the whole society. It includes three parts with a total of fifteen articles: 1. Author publication ethics (Academic papers are an important part of scientific research; Academic misconduct is a cancer of academic papers; Authors are the main contributors to academic papers; Author signature reflects authors' intellectual property rights and academic contributions; Academic papers should attach great importance to intellectual property rights and information security; Standardized citation of references is an important manifestation of academic norms; Great importance should be attached to the standardization of research data and management; Establish error correction and academic self-purification mechanisms). 2. Peer review expert publication ethics (Peer review is an important quality control mechanism for papers; Review experts should comply with relevant requirements for paper review; Review experts should strictly follow relevant ethical guidelines and codes of conduct). 3. Editor publication ethics (Editors should become guardians of academic paper quality; Editors should play a monitoring role in academic ethics construction; Editors should become the last barrier to curb academic misconduct; Implement "zero tolerance" for academic misconduct).

Full text available at: http://www.lis.ac.cn/CN/column/column291.shtml

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