A Temporal Perspective on Academic Contributions of Major Countries in Library and Information Science: Postprint

Authors: Liu Hong, Li Yu

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

Abstract

[Purpose/Significance] To reveal the academic contributions and evolutionary characteristics of major countries in the field of library and information science since 2002 across dimensions including total publication output, citations, research topics, and academic collaboration, thereby providing references for discipline construction in LIS and for researchers tracking research topics and frontiers.

[Method/Process] Taking 86 core journals in the field of library and information science indexed in the WoS database over the past 15 years as the research object, this study applies research methods including bibliometrics, scientific knowledge mapping, and econometric models, and utilizes data analysis tools such as TDA, Ucinet, and VOSviewer to conduct an in-depth analysis of the academic contributions of major countries in LIS from three perspectives: total output characteristics, citation characteristics, and thematic characteristics of academic papers, and constructs an econometric model to examine the influence mechanisms of spatial distance and academic capacity on academic collaboration among major countries.

[Results/Conclusions] The United States demonstrates strong research strength in the field of library and information science, China exhibits the fastest growth in academic contributions, while the Netherlands, Finland, and Belgium perform prominently on the normalized citation ratio indicator; research topics are mainly concentrated in eight thematic categories: library, information retrieval, Internet, knowledge management, social networks, metrics, user studies, and research methods, with the United Kingdom, Spain, China, and the United States having the most extensive research topics; both spatial distance and academic capacity exert significant influence on international academic collaboration.

Full Text

Abstract

[**Purpose/Significance**] This study reveals the academic contributions and evolutionary characteristics of major countries in the field of library and information science (LIS) since 2002 across multiple dimensions, including total scholarly output, citations, research themes, and academic collaboration. The findings provide valuable references for discipline construction in LIS and for researchers tracking research topics and frontiers. [**Method/Process**] Focusing on 86 core journals in LIS indexed in the Web of Science (WoS) database over the past 15 years, this research employs bibliometric analysis, scientific knowledge mapping, and econometric models, utilizing analytical tools such as TDA, Ucinet, and VOSviewer. The study conducts an in-depth analysis of major countries' academic contributions from three perspectives: publication volume characteristics, citation features, and thematic features, while constructing an econometric model to examine the influence mechanisms of spatial distance and academic capacity on academic collaboration among these countries. [**Result/Conclusion**] The United States demonstrates strong research strength in LIS, while China shows the fastest growth in academic contributions. The Netherlands, Finland, and Belgium excel in normalized citation rate indicators. Research themes concentrate in eight categories: libraries, information retrieval, internet, knowledge management, social networks, metrology, user studies, and research methods, with the United Kingdom, Spain, China, and the United States covering the broadest range of topics. Both spatial distance and academic capacity significantly influence international academic collaboration.

2 Framework Design and Implementation Mechanism

2.1 Framework Design

To comprehensively analyze the academic contributions of major countries in LIS, this study constructs a research framework as shown in Figure 1 [Figure 1: see original paper].

First, query construction and data collection. This study examines 86 core journals in LIS from the WoS Core Collection (SCI, SSCI, A&HCI) between 2002 and 2016, limiting document types to "Article" and using the advanced search expression "SU=library and information science." This retrieval yielded 47,758 records from 154 countries/regions, designated as dataset D1. Simultaneously, based on author affiliation fields without distinguishing author order (i.e., counting each appearance of a country in a paper without multiple counting or ranking), we extracted 39,647 records from 19 major countries accounting for over 1% of total publications, designated as dataset D2. We also obtained 913,877 references from these 19 countries during 2002–2016, designated as dataset D3. Using WoS's "Create Citation Report" function, we retrieved citation data for these 19 countries' publications from 1999–2013 (with

citations in 2002–2016 from all 154 countries/regions), totaling 81,970 records, designated as dataset D4. The retrieval was conducted from January 10–13, 2018.

Second, data cleaning. The cleaning process focused on country and keyword fields. (1) Country field cleaning: Merging different regional entries for the same country (e.g., combining mainland China with Hong Kong and Macau, and merging England, Scotland, and Northern Ireland). (2) Keyword field cleaning: Comparative analysis revealed that author keywords in WoS better represent research themes than system-assigned keywords, indicating high academic standards among LIS scholars in literature formatting. To clarify knowledge themes, we standardized author keywords by merging synonyms, splitting compound terms, and removing invalid keywords. For example, "information literacy" in WoS includes four variants: information literacy, information literacies, information alliteracy, and literacy:information, which required standardization. This process reduced the 72,086 keywords in dataset D1 to 44,406.

Third, high-frequency term selection and co-occurrence matrix construction. We selected 19 countries with over 1% publication share as major countries. Following J.C. Donohue's algorithm based on Zipf's second law for determining high-frequency keyword thresholds, we calculated the number of high-frequency keywords using Formula 1, where I represents the number of high-frequency keywords and I_1 represents keywords appearing only once.

$$8I_1\sqrt{+1}-1$$

Fourth, data analysis. This includes visual analysis of research themes and empirical analysis of factors influencing academic collaboration.

2.2 Research Methods and Implementation Mechanism

Using the four standardized datasets (D1, D2, D3, D4), this study conducts visual analysis of major countries' academic contributions in LIS from four aspects.

First, based on datasets D2 and D3, we employ time-series analysis to compare annual publication volume and citation patterns of major countries, examining overall trends in scholarly output.

Second, based on dataset D4, we calculate the annual normalized citation ratio (RoC) for major countries using L. Bornmann et al.'s method (Formula 2) to compare citation impact across countries. In Formula 2, R , represents country a's normalized citation ratio in year t; P , is the number of publications by country a in year t; C , $_3$ represents citations received in year $t+3$ by these publications from country j; M is the number of countries publishing in year t; and N is the number of countries citing these publications in year $t+3$.

$$\frac{\sum_{j=1}^{N} C_{aj,t+3}}{\sum_{i=1}^{M}\sum_{j=1}^{N} C_{ij,t+3}} \Big/ \frac{P_{a,t}}{\sum_{i=1}^{M} P_{i,t}}$$

Third, based on datasets D1 and D2, we construct thematic knowledge maps for LIS. Using NetDraw integrated in Ucinet, we create thematic knowledge maps and country-theme two-mode knowledge maps for different time windows. VOSviewer is used to generate academic collaboration networks among major countries across time windows, capturing the evolution of research themes and applying content analysis to summarize thematic categories and their distribution across countries.

Fourth, based on dataset D2, we explore factors influencing academic collaboration among major countries. We construct an econometric model incorporating spatial distance and academic capacity variables to examine mechanisms affecting collaboration, with empirical analysis conducted using Stata.

3 Analysis of Total Publication Output Characteristics

3.1 Temporal Analysis of Total Publication Output

Based on bibliometric analysis methods, countries' publication volumes in LIS directly reflect their research capacity and academic status [10] (see Table 1 and Figure 2 [Figure 2: see original paper]).

Using SPSS 22.0 for curve fitting, we found that a linear function $y = 0.046x^2$ - *181412.72 (*$R^{2*} = 0.98$) best fits the temporal evolution of total publications from major countries. The United States and the United Kingdom rank first and second respectively in total publications in the WoS Core Collection, with the U.S. far outpacing other countries, accounting for 41.86% of total output from all major countries. Notably, China and Spain have demonstrated rapid growth, with China's publications increasing sharply since 2004 and surpassing the United Kingdom in 2012, while Spain has consistently ranked fourth since 2008. Canada, Australia, Germany, the Netherlands, South Korea, and Brazil constitute the second tier of publication output.

3.2 Component Clustering Based on Publication Output

To better reflect temporal variations in publication output, we introduce the coefficient of variation (CV), the ratio of standard deviation to mean. A higher CV indicates greater fluctuation in a country's annual output over the 15-year period. The mean publication volume and mean CV for major countries are 158 and 0.41, respectively. Drawing on the Boston Matrix concept, we use these means as axes to construct a two-dimensional matrix, dividing the 19 countries into four components (see Figure 3 [Figure 3: see original paper]).

Component K1 (high market share but below-average growth rate) comprises the United States, United Kingdom, and Canada—traditional research power-

houses in LIS. **Component K2** (above-average market share and growth rate) includes China and Spain—emerging research powers. **Component K3** (low market share but above-average growth rate) consists of Brazil, South Korea, Italy, South Africa, and seven other countries with research potential. **Component K4** (below-average market share and growth rate but maintaining upward momentum) includes Finland, the Netherlands, Singapore, and Belgium—countries with slowing research growth.

4 Analysis of Citation Characteristics of Publication Output

Citation analysis objectively evaluates the recognition and attention received by countries' research achievements and measures their influence in international knowledge exchange and cooperation [11]. This study employs two indicators—citation volume and normalized citation ratio (RoC)—for temporal analysis.

4.1 Temporal Analysis of Citation Volume

Analysis results for major countries' citation volumes are presented in Table 2 and Figure 4 [Figure 4: see original paper]. The United States shows citation growth far exceeding other countries, with a mean citation share of 51.74%. The United Kingdom, Canada, and China rank second, third, and fourth respectively. Combined with Figure 5 [Figure 5: see original paper], which compares publication and citation rankings, the U.S. and U.K. dominate both metrics, reflecting long-term government and funding agency support for LIS and accumulated academic achievements. Canada, as a K1 research powerhouse alongside the U.S. and U.K., also demonstrates substantial impact, ranking third in citation influence. China shows remarkable growth in both publications and citations over the past 15 years, ranking third and fourth respectively, indicating strong research momentum despite a slightly lower citation ranking. Conversely, Brazil, Spain, and Italy exhibit citation rankings significantly lower than their publication volumes, suggesting a need to improve publication quality. Denmark, Finland, and Singapore, while ranking lower in output, achieve citation performance exceeding their publication rankings, representing a "small but excellent" model. Australia, Germany, South Korea, and Sweden show consistent rankings between output and citation metrics.

4.2 Temporal Analysis of Normalized Citation Ratio

Applying Formula 2, we calculated the annual RoC for major countries from 1999–2013 and computed mean RoC values (see Table 3), classifying countries into three expectation-based subgroups. The evolution of each subgroup's RoC over time is shown in Figure 6 [Figure 6: see original paper].

(1) Exceeding expectations (RoC $\geq$ 1.2): Eleven countries exceed expectations, with the Netherlands, Finland, and Belgium ranking top three. Despite

modest publication volumes, these countries achieve high citation levels internationally. The United States and Canada also achieve RoC > 1.2, indicating exceptional quality alongside high productivity, suggesting they will maintain their leadership positions. Notably, China's RoC reaches 1.53, demonstrating strong research strength in both quantity and quality.

(2) Meeting expectations (0.8 ≤ RoC < 1.2): The United Kingdom, Italy, Spain, and Germany meet expectations, with the U.K. achieving an RoC of 1.14, maintaining high-quality output amid rapid technological development and paradigm shifts.

(3) Below expectations (RoC < 0.8): Brazil, France, India, and South Africa fall below expectations, lagging in both publication quantity and quality.

5 Thematic Knowledge Mapping Analysis of Publication Output

Keywords, as distillations of research content, reflect study themes and serve as important indicators in bibliometric research [12]. Constructing thematic knowledge maps from keyword co-occurrence matrices enables visual interpretation of research content and better understanding of knowledge development trajectories and frontiers [13]. This study analyzes LIS thematic knowledge maps from temporal and country dimensions.

5.1 Evolution of Thematic Knowledge Mapping in Different Time Windows

5.1.1 Thematic Knowledge Network Mapping. Using dataset D1, we divided the 15-year period into three equal time windows: T1 (2002–2006), T2 (2007–2011), and T3 (2012–2016). Applying Donohue's high-frequency threshold algorithm (Formula 1), we identified 120 high-frequency keywords for T1 (4,515 occurrences, 56.15% of total), 172 for T2 (10,583 occurrences, 59.03%), and 219 for T3 (17,902 occurrences, 63.23%). Using Thomson Data Analyzer to construct co-occurrence matrices and NetDraw for visualization, we created thematic knowledge maps for each window (see Figure 7 [Figure 7: see original paper]). The network structures show increasing density and more frequent co-occurrence behavior among high-frequency keywords across time windows.

5.1.2 Thematic Category Analysis. Thematic analysis reveals eight major categories: libraries, information retrieval, internet, knowledge management, social networks, metrology, user studies, and research methods.

- **T1 period:** Themes include "internet" (linked to e-commerce, ICT, digital divide), "libraries" (linked to interlibrary loan, document delivery, e-journals), and "information retrieval" (linked to search engines, classification, databases).
- **T2 period:** Themes expand to include "internet" (evolving from Web

1.0 to Web 2.0), "information retrieval" (linked to semantic web, digital libraries), "user studies" (focused on information needs, behavior, literacy), "knowledge management" (linked to knowledge sharing, transfer, creation), "bibliometrics" (measuring research performance and collaboration), "social networks" (studying Twitter, Blogs, Facebook with sentiment analysis and privacy issues), and "qualitative research methods" (frequently using interviews, case studies, and grounded theory in healthcare contexts).

- **T3 period:** Themes include "information retrieval" (shifting focus to information needs, behavior, literacy), "academic libraries" (linked to open access, social networks, information literacy), "knowledge management" (increasing emphasis on intellectual property and social network-based disaster management), "bibliometrics," "social networks," and "qualitative research methods."

5.1.3 Highly Cited Paper Analysis. Content analysis of highly cited papers across time windows and thematic categories identifies ESI highly cited papers from the past 15 years (see Table 4). These papers focus on social networks, user behavior, big data, and information systems, primarily published in *MIS Quarterly*, *Journal of Computer-Mediated Communication*, and *JASIST*.

5.2 Two-mode Knowledge Mapping Analysis of Country-Theme

5.2.1 Two-mode Knowledge Mapping. Two-mode networks analyze relationships between two different node types, specifically affiliation networks [25]. Our network comprises countries and keywords as nodes. Since dataset D2 counted all authors' contributions (potentially double-counting multi-country papers), it provides comprehensive assessment of countries' contributions to different themes. We selected 114 keywords with frequency $ 100, *constructinga*19×$114 two-mode matrix. To manage complexity, we retained only keywords in the top quartile of association strength for each country and removed keywords with association strength <25, resulting in a final matrix with 19 country nodes and 65 keyword nodes, visualized in Figure 8 [Figure 8: see original paper].

5.2.2 Two-mode Network Centrality Analysis. Centrality analysis results for country and keyword nodes are shown in Table 5 . Social networks, information retrieval, and knowledge management represent the hottest research topics with the strongest research forces. The United Kingdom, Spain, China, and the United States exhibit the broadest research themes. The U.K. excels in library science topics (document delivery, public libraries, e-books). Spain leads in metrology topics (webometrics, semantic web, research evaluation). China shows strong performance in text mining, data mining, H-index, user satisfaction, and technology acceptance models. The United States demonstrates depth in digital divide, digital libraries, trust/privacy issues, and knowledge sharing/decision support.

6 Analysis of Factors Influencing Academic Collaboration Among Major Countries

6.1 Factor Discovery Based on Country Collaboration Knowledge Mapping

As scientific collaboration deepens, the international academic collaboration landscape in LIS evolves accordingly. We constructed collaboration networks for three time windows among 19 major countries to explore collaboration mechanisms. In Figure 9 [Figure 9: see original paper], node sizes represent publication volumes, line thickness indicates collaboration intensity, and node colors represent clusters identified by VOSviewer.

Figure 9 shows increasing collaboration frequency, with network density reaching 0.93 in T3. The United States, United Kingdom, China, and Canada consistently rank top in centrality, maintaining solid collaborative relationships. Other countries show relatively stable collaboration partners over time. Setting a collaboration threshold of 10 reveals clear regional patterns: Sweden collaborates only with European neighbors (Netherlands, Germany, Finland), while Singapore collaborates only with Asian neighbor South Korea. Studies by Acosta et al. [26] and Parreira et al. [27] confirm spatial distance as a key factor, as collaboration costs increase with distance.

6.2 Empirical Analysis of the Academic Collaboration Influence Factor Model

Building on Zhang and Ouyang's [28] gravity model for academic collaboration, we construct our model (Formula 3), where H , represents collaborative publications between countries i and j in period t; P_1 , and P_2 , represent their respective publication outputs (academic capacity); D is spatial distance (calculated from capital city coordinates); and , is the error term.

$$H_{ij,t} = \alpha_0 + \alpha_1 \ln P_{1i,t} + \alpha_2 \ln P_{2j,t} + \alpha_3 \ln D_{ij} + \varepsilon_{ij,t}$$

The 19 countries include 4 Asian, 10 European, 2 North American, and 1 each from South America, Oceania, and Africa. We analyze collaboration models by continent (Asia, Europe, North America) and time window (T1, T2, T3). Since H , is a count variable, negative binomial regression is appropriate. Empirical results using Stata 14.0 are shown in Table 6 .

Results indicate that both countries' research capacities significantly and positively influence collaboration, while spatial distance has a significant negative effect, though academic capacity's impact is stronger. Improving publication quality is thus an effective way to overcome spatial barriers and promote international exchange.

(1) Impact of academic capacity: Asian countries' collaborations are more strongly influenced by academic capacity. Across continents, collaborators' aca-

demic capacity has greater impact than one's own capacity, and this influence strengthens over time.

(2) Impact of spatial distance: Asian countries (China, South Korea, Singapore, India) are most affected by spatial distance, with frequent intra-regional collaboration but limited intercontinental exchange. European countries also show regional tendencies, while North American countries (U.S., Canada) are less constrained by distance, demonstrating better international perspectives. Over time, regionalization of academic exchange has intensified, contrary to assumptions about increasing internationalization, possibly due to political, economic, and cultural cooperation within regions and facilitation by organizations like the EU and APEC.

7 Conclusions and Discussion

This study collected 15 years of publication and citation data for 19 major LIS countries, constructing co-word matrices and scientific knowledge maps to analyze contributions from publication, citation, and thematic perspectives, while revealing mechanisms influencing academic collaboration.

The United States demonstrates formidable research strength and outstanding academic contributions in LIS. China shows the fastest growth, ranking second in publication volume in recent years. The United Kingdom, Canada, Spain, and Australia also rank high across dimensions, while the Netherlands, Finland, and Belgium exceed expectations in normalized citation rates. Research themes concentrate in eight categories, with the U.K., Spain, China, and the U.S. covering the broadest ranges. Spatial distance and academic capacity significantly affect collaboration, with both factors' influence increasing over time. Among the three continents, Asian countries' collaborations are most affected by these factors.

Limitations: This study comprehensively reveals international academic contributions in LIS at the publication level but does not examine countries' participation in international conferences, academic institutions, or journal editorial boards, nor does it assess the social impact of their research. These aspects represent directions for future work.

References

[1] Duan Xiane, Zhang Xiaojuan. Analysis of China's international publications in library and information science (2011–2015) [J]. *Library Tribune*, 2017, 37(5): 64–73.

[2] Shi Xiaoxiao, Zhang Qingpu. Research status and development trends of foreign disruptive innovation based on co-word analysis [J]. *Journal of the China Society for Scientific and Technical Information*, 2017, 36(7): 748–759.

[3] Li Xin, Li Xuhui, Lu Wei, et al. Mining hot fields in library and information

science driven by big data: An empirical perspective based on WoS bibliographic data [J]. *Library Tribune*, 2017, 37(4): 49–57.

[4] Tan Chunhui, Zeng Juan, Qiu Junping. Analysis of domestic intelligence science research trends during the 12th Five-Year Plan period based on CSSCI [J]. *Journal of the China Society for Scientific and Technical Information*, 2017, 36(7): 734–747.

[5] Zhao Rongying, Wei Xuqiu. Evolution of international cooperation and exchange features based on comparative static analysis: A case study of Chinese library and information science [J]. *Library*, 2017(8): 27–31.

[6] Tong Shouchuan, Li Jiang, Li Dong. Current status and trends of internationalization development of Chinese library and information science [J]. *Library and Information Service*, 2017(6): 24–34.

[7] Hou Jianhua, Yang Xiucai, Zhou Lijuan. Analysis of frontier topics and evolutionary trends in international library and information science research [J]. *Library and Information Service*, 2016, 60(13): 82–90.

[8] DONOHUE JC. *Understanding Scientific Literatures: A Bibliometric Approach* [M]. Cambridge: MIT Press, 1973: 129.

[9] BORGMANN L, WAGNER C, LEYDESDORFF L. The geography of references in elite articles: Which countries contribute to the archives of knowledge? [J]. *PloS One*, 2018, 13(3): e0194805.

[10] VAN ECK NJ, WALTMAN L. Software survey: VOSviewer, a computer program for bibliometric mapping [J]. *Scientometrics*, 2010, 84(2): 523–538.

[11] He Wen, Liu Ying, Yang Hongmei. Research on international co-authorship and contributions of highly cited papers in LIS: A bibliometric perspective [J]. *Library and Information Service*, 2017(2): 83–91.

[12] Ke Ping, Gong Ping. Analysis of evolution path and hot fields in digital humanities research [J]. *Journal of Library Science in China*, 2016, 42(6): 13–30.

[13] Hu Zewen, Lian Tonghui, Cao Ling. Scientific knowledge mapping analysis of network resource aggregation and navigation research [J]. *Information Science*, 2017, 35(7): 9–15.

[14] HERFORT B, BRENNING A, ZIPF A. A geographic approach for combining social media and authoritative data toward identifying useful information for disaster management [J]. *International Journal of Geographical Information Science*, 2015, 29(4): 667–689.

[15] ELLISON NB. Social network sites: Definition, history, and scholarship [J]. *Journal of Computer-Mediated Communication*, 2007, 13(1): 210–230.

[16] PETTER S, STRAUB D, RAI A. Specifying formative constructs in information systems research [J]. *MIS Quarterly*, 2007, 31(4): 623–656.

[17] VENKATESH V, THONG JYL, XU X. Consumer acceptance and use of information technology: Extending the unified theory of acceptance and use of technology [J]. *MIS Quarterly*, 2012, 36(1): 157–178.

[18] LIBEN-NOWELL D, KLEINBERG J. The link prediction problem for social networks [J]. *Journal of the Association for Information Science and Technology*, 2007, 58(7): 1019–1031.

[19] WETZELS M, ODEKERKEN-SCHRÖDER G, OPPEN CV. Using PLS path modeling for assessing hierarchical construct models: Guidelines and empirical illustration [J]. *MIS Quarterly*, 2009, 33(3): 177–195.

[20] CHEN H, CHIANG RHL, STOREY VC. Business intelligence and analytics: From big data to big impact [J]. *MIS Quarterly*, 2012, 36(4): 1165–1188.

[21] PFEFFERS K, TUUNANEN T, ROTHENBERGER MA, et al. A design science research methodology for information systems research [J]. *Journal of Management Information Systems*, 2007, 24(3): 45–77.

[22] VALENZUELA S, PARK N, KEE KF. Is there social capital in a social network site? Facebook use and college students' life satisfaction, trust, and participation [J]. *Journal of Computer-Mediated Communication*, 2009, 14(4): 875–901.

[23] JANSEN BJ, ZHANG M, SOBEL K, et al. Twitter power: Tweets as electronic word of mouth [J]. *Journal of the Association for Information Science and Technology*, 2009, 60(11): 2169–2188.

[24] HSU CL, LIN JCC. Acceptance of blog usage: The roles of technology acceptance, social influence, and knowledge sharing motivation [J]. *Information & Management*, 2008, 45(1): 65–74.

[25] BORGATTI SP, MEHRA A, BRASS DJ, et al. Network analysis in the social sciences [J]. *Science*, 2009, 324(5926): 892–895.

[26] ACOSTA M, CORONADO D, FERRANDIZ E, et al. The geography of university scientific production in Europe: An exploration in the field of food science and technology [J]. *Scientometrics*, 2017, 112(1): 215–240.

[27] PARREIRA MR, MACHADO KB, LOGARES R, et al. The roles of geographic distance and socioeconomic factors on international collaboration among ecologists [J]. *Scientometrics*, 2017, 113(3): 1539–1550.

[28] Zhang Cui, Ouyang Dongping. Research on factors influencing China's international academic cooperation under the Belt and Road Initiative: An empirical analysis based on co-authored papers in the Web of Science [J]. *Journal of International Trade*, 2017(4): 74–82.

Author Contributions: Liu Hong: Framework design, manuscript writing and revision; Li Yu: Data processing.

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