Scientometrics (2015) 105:2071–2087
DOI 10.1007/s11192-015-1762-8

Evolution of research subjects in library and information science based on keyword, bibliographical coupling, and co-citation analyses

Yu-Wei Chang[1] · Mu-Hsuan Huang[1] · Chiao-Wen Lin[1]

Received: 14 May 2015 / Published online: 26 September 2015
© Akadémiai Kiadó, Budapest, Hungary 2015

Abstract This study involved using three methods, namely keyword, bibliographic coupling, and co-citation analyses, for tracking the changes of research subjects in library and information science (LIS) during 4 periods (5 years each) between 1995 and 2014. We examined 580 highly cited LIS articles, and the results revealed that the two subjects "information seeking (IS) and information retrieval (IR)" and "bibliometrics" appeared in all 4 phases. However, a decreasing trend was observed in the percentage of articles related to IS and IR, whereas an increasing trend was identified in the percentage of articles focusing on bibliometrics. Particularly, in the 3rd phase (2005–2009), the proportion of articles on bibliometrics exceeded 80 %, indicating that bibliometrics became predominant. Combining various methods to explore research trends in certain disciplines facilitates a deeper understanding for researchers of the development of disciplines.

Keywords Library and information science · Keyword analysis · Bibliographical coupling · Co-citation

Introduction

Researchers communicate their research findings by presenting information and using formal channels. Among such channels, peer-reviewed journals are regarded as pivotal in most disciplines. Researchers typically access articles to track certain research subjects in their own disciplines. Therefore, subject analysis of journal articles in specific disciplines enables researchers to determine the range of research topics and identify new topics. Trend analysis in research subjects assists researchers in planning their research direction

✉ Mu-Hsuan Huang
mhhuang@ntu.edu.tw

[1] Department of Library and Information Science, National Taiwan University, Taipei, Taiwan

and predicting research trends. Thus, trends in research subjects have been of considerable concern for academics.

Previous studies have used various methods to present the trends of research topics for specific disciplines and visualize the intellectual structure of disciplines (Jarneving 2005; White and Griffith 1981; White and McCain 1998). The main advantage of bibliometric analysis is that it assists researchers in observing historical changes in the characteristics of publications on the basis of analyses conducted using ample bibliographic data. The subject relationship between the citing and cited publications, typically investigated through citation analysis, has been widely used to analyze bibliographic data in numerous disciplines. In addition to direct citation, both bibliographical coupling and co-citation, two additional types of citation analysis, have been used to identify subfields or research subjects in various disciplines. However, related studies have shown that direct citation has been used the most and bibliographical coupling has been used the least.

Both relationships of bibliographical coupling and co-citation between two documents are created through a third document. When two documents are cited by a third document, a co-citation relationship exists between them. The more documents cite a pair of documents, the higher is the co-citation strength of the two documents (Small 1973). Similarly, when two documents cite the same document, they are bibliographically coupled (Kessler 1963). The higher the number of common documents cited by two documents, the higher is the bibliographic coupling strength of the two documents. A high value in both bibliographic coupling strength and co-citation strength indicates a similar subject relationship between the two documents.

Although both bibliographical coupling and co-citation analyses can be used to elucidate the intellectual structure of disciplines, researchers have preferred co-citation analysis. A few researchers have claimed that co-citation analysis is superior in displaying disciplinary structures to bibliographical coupling analysis (Bichteler and Eaton 1980; Sharabchiev 1989). However, bibliographic coupling has been proven to be an effective method for identifying changes in research topics (Zhao and Strotmann 2008, 2014). The number of studies using bibliographic coupling analysis has increased (Boyack and Klavans 2010; Ma 2012; Zhao and Strotmann 2014; Yan and Ding 2012). In addition, researchers are increasingly investigating intellectual structures of disciplines by using various bibliometric techniques and comparing the differences and similarities of results generated by various bibliometric techniques (Åström 2002; Boyack and Klavans 2010; Qiu, Dong, and Yu 2014; Yan and Ding 2012; Zhao and Strotmann 2014).

Combining various bibliometric analyses can reveal details of research subjects for specific disciplines, because each bibliometric analysis exhibits certain advantages. In this study, three bibliometric types of analyses (keyword analysis, bibliographic coupling, and co-citation analysis) were used to explore research subjects in library and information science (LIS); the analyses differ from those used in previous studies. In addition, the data that were analyzed in this study differ from those of previous studies. Although a comparison of results generated through various methods seems difficult, numerous characteristics of research subjects embedded in LIS can be revealed, because each method involves a unique approach. Because of constant changes in research subjects, investigating the trends in research subjects of various disciplines remains essential. The development of disciplines necessarily involves the growth and decline of research subjects and the emergence of new research topics; this process is integral to knowledge development. Therefore, researchers tend to monitor research subjects. This study focused on the changes in research subjects in LIS during four successive 5-year periods between 1995 and 2014. The differences among the results obtained using three distinct bibliometric

methods are anticipated to facilitate a deeper understanding of the evolution of LIS subjects. The primary focus of this study concerned the common research subjects obtained by using three distinct bibliometric methods.

Literature review

Studies on changes in LIS research topics can be divided into three groups according to the research methods used, including content analysis, bibliographic analysis, and a combination of various methods. Studies using content analysis have devised or adopted specific classification schemes of other researchers to classify the research subjects of LIS articles (Chang 2004; Chen and Liang 2004; Järvelin and Vakkari 1993; Koufogiannakis et al. 2004; Shi 2002; Tsay and Hsu 2009). Compared with content analysis, bibliometric analysis comprises numerous techniques, and researchers have investigated the differences among analyses conducted using such techniques.

Various bibliometric techniques involve researchers using keywords provided by authors or database providers to examine the research topics of articles. For example, citation index databases of Web of Science include two types of keywords: Author Keywords, generated by authors, and Keywords Plus, generated using the most frequently occurring words and phrases in the titles of documents cited by authors (Garfield 1992–1993). These two types of keywords have been used in numerous studies (Chiu and Ho 2007; Tsai et al. 2008; Yi et al. 2008). Regarding LIS, Tsay and Lai (2007) used descriptors obtained from the database of Library and Information Science Abstracts (LISA) to investigate the research topics in information science. They identified that professional education and information retrieval were two prominent subjects.

Both co-citation and bibliographical coupling represent co-occurrence networks. In addition to documents, authors and journals have been used as the unit of analysis for identifying co-occurrence networks (Ding et al. 2000; Hu et al. 2011; Ma et al. 2009; Qiu et al. 2014; White and Griffith 1981; White and McCain 1998; Zhao and Strotmann 2014). After document co-citation analysis was introduced in the 1970s, authors became a widely used unit of analysis for investigating co-citation networks. De-Moya-Anegon et al. (1998) adopted author co-citation analysis to investigate the subject structure and research fronts of Spanish LIS publications between 1985 and 1994. "bibliometrics and informetrics" and "librarianship" were identified as the two main subjects. Ma et al. (2009) used author co-citation analysis to determine the intellectual structure of information science in China.

Bibliographical coupling analysis has seldom been used to explore the intellectual structure of disciplines. However, using this approach has been increasingly observed. Zhao and Strotmann (2008) used author bibliographic coupling to present the intellectual structure of information science and reported that bibliographic coupling and co-citation supplement each other. Ma (2012) used author bibliographic coupling to reveal the intellectual structure of LIS in China. Qiu et al. (2014) compared the following five types of author co-occurrence networks on the basis of 30 LIS journals: co-authorship, author co-citation, author bibliographic coupling, words-based author coupling, and journal-based author coupling.

Because each bibliometric analysis features unique characteristics, researchers have begun combining various methods to create a comprehensive map of the intellectual structure of science. Åström (2002) used the following three methods to investigate research subjects on the basis of nine LIS journals articles between 1998 and 2000:

keyword co-occurrence, author co-citation, and a combination of keyword occurrence and author co-citation. The 52 most cited authors and the 47 most frequent occurring keywords collected from descriptors of the Educational Resources Information Center (ERIC) database were used to form a matrix. A comparison of cluster analyses generated from three methods (two methods for keyword co-occurrence and a combination of keyword co-occurrence and author co-citation) enabled generating the two similar clusters of information retrieval and bibliometrics.

Previous studies have revealed a slight increase in the number of studies combining various bibliometric methods to investigate the intellectual structure of LIS. However, most studies have emphasized the differences among the used bibliometric analyses. Few studies were longitudinal studies on the evolution of research subjects in LIS. Åström (2007) observed changes in LIS research fronts using document co-citation analysis in three 5-year periods (1990–2004). Zhao and Strotmann (2008) compared the changes in research subjects of information science between 1996–2000 and 2001–2005 according to results generated from co-citation and bibliographic coupling analyses. To investigate the evolution of research subjects in LIS, we monitored the changes in research subjects of LIS during four successive 5-year periods (1995–1999, 2000–2004, 2005–2009, and 2010–2014). The results were strengthened by the combination of three bibliometric analyses, namely keywords, bibliographic coupling, and co-citation. The two focuses in this study were the changes in the research subjects in LIS and the differences in the research subjects in LIS resulting from the use of the three methods.

Methodology

In this study, we used keyword, bibliographic coupling, and co-citation analyses to investigate research subjects according to articles in LIS journals between 1995 and 2014. Ten LIS journals that satisfied three requirements were used. First, journals had to be listed in the subject category of information science and library science according to the 2008 version of the Journal Citation Reports. If a journal changed its title during the study period (1995–2014), then the previous and current titles were regarded as the same journal. Second, several journals focusing on management and computer science were included in the subject category information science and library science. To ensure that the selected journals were LIS-oriented, we defined LIS journals as those that were indexed simultaneously in three LIS databases of LISA; Library Literature and Information Science; and Library, Information Science and Technology Abstracts. Third, only ten journals (Table 1) with the highest impact factors were analyzed.

Each journal title was used to retrieve research articles published during 1995–2014. Only research articles coded as "articles" were analyzed for this study. Because we identified 10,864 articles, the sample articles were obtained using the following process. First, the 10,864 articles were divided into four groups. Each of the four groups contained articles published in a period of five successive years: 1995–1999, 2000–2004, 2005–2009, and 2010–2014. Second, the top 5 % of highly cited articles published in the four periods were identified individually to ensure that the primary research subjects could be revealed. Because articles need time to accumulate the number of citations they receive, the top 5 % of highly cited articles published between 2010 and 2014 were selected annually for the inclusion of current articles. Third, to ensure that the top 5 % of highly cited articles were LIS-oriented, the research topics were examined according to the titles, abstracts, and

Table 1 LIS journals used in this study

No.	Journal title	No. of articles
1	Scientometrics	2516
2	Journal of the Association for Information Science and Technology (Previous title: Journal of the American Society for Information Science; Journal of the American Society for Information Science and Technology)	2546
3	Government Information Quarterly	710
4	Information Processing and Management	1188
5	Journal of Documentation	643
6	Journal of the Medical Library Association (Previous title: Bulletin of the Medical Library Association)	854
7	Journal of Information Science	866
8	Library and Information Science Research	451
9	College and Research Libraries	658
10	Information Technology and Libraries	432
	Total	10,864

keywords. When a non-LIS-oriented article was excluded, another LIS-oriented article was added. A total of 580 articles represented in nine LIS journals were selected. The journal *Information Technology and Libraries* was excluded from the journal list in Table 1 because none of the selected articles were published in this journal.

To enable keyword analysis, the keywords of articles were added from the descriptors in the database of LISA. Each article indexed in LISA was allotted one or more descriptors. Each descriptor was counted once, and then the ranking of keywords was generated. Because 81 articles were not indexed in LISA, their descriptors were created after we referred to the form of descriptors of LISA.

Regarding the bibliographic coupling analysis, four bibliographic coupling matrixes representing the subject relationship between two documents had to be formed first and then converted into Pearson correlation matrixes. All pairs of articles were examined to determine whether a pair of articles cited the same documents according to references listed in the pair of articles. The number of documents that two articles cited together represented the degree of bibliographical coupling exhibited. The pairs of articles with a value of 0 or 1 for the degree of bibliographical coupling were excluded because of a weak subject relationship. The four matrixes of bibliographical coupling were converted into Pearson correlation matrixes using UCINET for windows 6.165. The Pearson correlation coefficient, a measure used in similarity metrics, measures the strength and direction of a linear relationship between two documents. The value of the correlation coefficient ranges between -1 and $+1$. When two documents exhibit a strong positive correlation, the value of the correlation coefficient is near $+1$. An inverse relationship is described as a negative correlation. The correlations serve as criteria with which to group documents for cluster analysis.

Similar processes were performed for co-citation analysis. We constructed four document co-citation matrixes by examining the co-citation strength of a pair of articles. When two articles were cited together by another article, the two articles exhibited a co-citation strength of 1. The level of similarity between two articles was represented using the value of the correlation coefficient.

The formations of clusters generated on the basis of bibliographic coupling and co-citation analyses varied with various values of the selected correlation coefficients. To compare the clusters formed in the four phases, we chose a similar correlation coefficient ranging from 0.500 to 0.600 after having tested various values. The specific value of the correlation coefficient for each time phase was determined according to the clustered display of articles using various correlation coefficients. In addition, hierarchical cluster analysis was used according to the single linkage rule; that is, the distance between two clusters is determined according to the distance of the two closest documents in the two clusters. Although this linkage rule may group two documents with less similarity in the same cluster and cause clusters with inconsistent sizes, it is more appropriate than other linkage rules that generate too many clusters.

Table 2 Fifteen keywords with the highest frequencies in the four 5-year phases

1995–1999		2000–2004		2005–2009		2010–2014	
Keyword	% of articles	Keyword	% of articles	Keyword	% of articles	Keyword	% of articles
Searching	4.9	WWW	6.3	Citation analysis	7.3	Scholarly publications	7.4
Online IR	3.9	Online IR	5.7	Bibliometrics	4.6	Bibliometrics	5.2
Research	3.6	Searching	4.9	Periodicals	4.6	Citation analysis	3.0
Citation analysis	3.6	Citation analysis	3.6	Scholarly publications	4.4	Social networks	2.2
Scientometrics	2.9	Research	2.8	Articles	4.1	Co-authorship	2.0
Information seeking	2.9	Science	2.0	Ranking	2.8	Methods	1.4
Evaluation	2.3	Periodicals	2.0	Academic publishing	2.8	Research	1.4
Periodicals	2.3	Articles	1.8	Science &Technology	2.5	Electronic government	1.3
WWW	2.3	Information seeking	1.8	Scientometics	2.5	Impact factor	1.3
Bibliometrics	2.1	Internet	1.8	Performance measure	2.2	Citations	1.1
Impact factor	1.8	Evaluation	1.6	Online IR	2.1	Ranking	1.1
Internet	1.3	Scientometrics	1.6	WWW	2.1	Universities	1.1
Online databases	1.3	Students	1.6	H-index	1.9	Citation index	1.0
Performance measure	1.3	Search engines	1.6	Impact factor	1.9	H-index	0.9
Children	1.0	Web sites	1.6	Evaluation	1.8	Evaluation	0.8
Information work	1.0	Links	1.4			Webmetrics	0.8
Science	1.0	Search strategies	1.4				
User needs	1.0	Impact factor	1.4				
	40.7		44.7		47.5		31.7

Results

Keyword analysis

Table 2 lists the 15 keywords with the highest frequencies over the four 5-year periods according to LISA descriptors of articles. The proportion of articles containing at least one of the top 15 keywords in each period ranged between 31.7 and 47.5 %. In the first period (1995–1999), "searching" (4.9 %) was ranked as the most used keyword, followed by "online information retrieval (online IR)" (3.9 %). "World Wide Web (WWW)" (6.3 %) and "online IR" (5.7 %) were the top two keywords during 2000–2004. "Citation analysis" (7.3 %) accounted for the highest proportion in the third period (2005–2009). In the last period (2010–2014), "scholarly publications" (7.4 %) was the most used keyword. Although various keyword rankings were generated in the four periods, no traditional library-science-oriented keywords are listed in Table 2. Most keywords are subjects related to information science. An additional possible factor explaining this finding is that most journals selected for this study were information science oriented.

Because numerous keywords in Table 2 have similar meanings and exhibit close relationships, the keywords were further divided into five broad subjects (Table 3). Each keyword was classified as one or two broad subjects. If a keyword was assigned to n broad subjects, then the number of articles in one broad subject was $1/n$ of the number of articles. Table 4 presents the percentage of articles containing keywords related to a specific broad subject.

Table 4 shows that "bibliometrics" dominated in the three periods, 1995–1999, 2005–2009, and 2010–2014. Although "application of Internet technology (AIT)" accounted for the highest proportion during 2000–2005, bibliometrics was a widely researched subject according to its proportion in four phases. Particularly, the percentage of "bibliometrics" was as high as 81.8 % during 2005–2009, indicating the percentages of the other four broad subjects to decrease sharply. An additional high fluctuation in the percentage was identified in "AIT," which exhibited an increase after the emergence of the WWW. However, the high proportion of "AIT" did not extend to the subsequent period, 2005–2009. The decreasing trend across the four periods was identified only through the percentage of articles related to "information behavior (IB)."

Table 3 List of mapping subject areas and keywords

Subject areas	Keywords
Application of Internet technology	Electronic government; Internet; Links; Search engines; Search strategies; Searching; Social networks; Web sites; WWW
Bibliometric	Academic publishing; Articles; Bibliometrics; Citation analysis; Citation index; Citations; Co-authorship; Evaluation; H-index; Impact factor; Performance measure; Periodical; Ranking; Scholarly publications; Scientometrics; Webmetrics
Information retrieval and system evaluation	Evaluation; Online databases; Online IR; Performance measure; Search strategies; Searching
Information behavior	Information seeking; User needs
Other	Children; Information work; Methods; Research; Science; Science and technology; Students; Universities

Table 4 Comparison of five subjects in the four phases using keyword analysis

1995–1999	% of articles	2000–2004	% of articles	2005–2009	% of articles	2010–2014	% of articles
Bibliometrics	35.7	AIT	35.4	Bibliometrics	81.8	Bibliometrics	74.2
IRSE	23.3	Bibliometrics	24.8	IRSE	8.6	Other	12.3
Other	16.6	IRSE	25.7	Other	5.2	AIT	10.9
AIT	15.0	Other	14.2	AIT	4.4	IRSE	2.6
IB	9.6	IB	4.0	IB	0.0	IB	0.0

IRSE Information retrieval and system evaluation, *AIT* application of internet technology, *IB* information behavior

Bibliographic coupling analysis

Tables 5 shows the results of the cluster analysis generated using articles with bibliographic coupling relationships in the four study periods. Nine clusters were present in 1995–1999. Cluster 1 ["information seeking (IS) and information retrieval (IR)"] comprised 34 articles and was considerably larger than any of the other eight clusters during 1995–1999. In 2000–2004, three out of ten clusters were prominent. The largest cluster related to IS and IR comprised 26 articles. However, the second and third largest clusters featured subjects related to bibliometrics. In 2005–2009, 15 clusters were identified and most articles were concentrated in the top two clusters. The largest cluster, "h-index," consisted of 54 articles, twice the size of the second largest cluster related to the visualization of the structure of science in terms of articles. In 2010–2014, "bibliometrics" was the largest cluster with 54 articles among 19 clusters. In addition, less than half of all clusters (9 clusters) consisted of only 2 or 3 articles. This differs from 3 other periods in which higher proportion of clusters with 2 or 3 articles appeared.

Various clusters were labeled with subjects having a similar or related meaning; therefore, clusters were further incorporated into broad clusters labeled with broad subjects. To compare the broad subjects generated using three methods, a classification list of broad subjects was adopted. We determined whether the list of broad subjects generated from keyword analysis corresponded with the results of the bibliographical coupling analysis. Because the keywords related to the two subjects of "information retrieval and system evaluation" and "information seeking behavior" were often present in the same articles, and an article was the basic unit analyzed for bibliographical coupling, the two subjects were incorporated into one broad subject, "IS and IR." Finally, the original list of five broad subjects was revised into a list comprising four broad subjects (Table 6).

Table 6 shows the numbers and percentages of articles with bibliographic coupling relationships in each broad subject during the four periods. The figures show that "IS and IR" and "bibliometrics" were the main research subjects in LIS between 1995 and 2004. However, "bibliometrics" was dominant in 2000–2004, 2005–2009, and 2010–2014. Among the four broad subjects, a prominent increasing trend was observed in "bibliometrics" between 1995 and 2009. A decreasing trend was identified in the percentages of articles on "IS and IR." "AIT", which did not emerge until 2000–2004, was the most recent and minor broad subject. An increasing trend also identified in "AIT".

Although a decreasing trend was observed in the percentage of articles on "IS and IR", a small cluster on seeking medical information was observed in the three periods between 1995 and 2009. Thus, researchers have been paying attention to seeking medical

Table 5 Cluster analysis based on articles with bibliographical coupling relationships during the four periods

Year	No. cluster	No. articles	Subject	Broad subject
1995–1999	1	34	Information seeking and information retrieval	IS and IR
	2	7	Citation analysis	Bibliometrics
	3	7	Visualization of science structure	Bibliometrics
	4	3	Doctors' information sources	IS and IR
	5	3	Bibliometric analysis of articles	Bibliometrics
	6	2	Attributes of graph files	IS and IR
	7	2	Electronic information seeking and usage of professional users	IS and IR
	8	2	Scientometrics	Bibliometrics
	9	2	Webmetrics	Bibliometrics
2000–2004	1	26	User-based information behavior	IS and IR
	2	18	Search engines study and webmetrics	Bibliometrics
	3	12	Citation analysis	Bibliometrics
	4	2	Information needs and information seeking in the field of clinical medicine	IS and IR
	5	2	Bibliometric analysis of articles	Bibliometrics
	6	2	Usage of search engines	AIT
	7	2	Internet application	AIT
	8	2	Webmetrics	Bibliometrics
	9	2	Co-citation analysis	Bibliometrics
	10	2	Interdisciplinary studies	Bibliometrics
2005–2009	1	54	H-index	Bibliometrics
	2	27	Visualization of science structure	Bibliometrics
	3	7	Information behavior	IS and IR
	4	5	Search engines	AIT
	5	4	E-government	Bibliometrics
	6	3	Bibliometric analysis of Astrophysics Data System	Bibliometrics
	7	3	Citation impact of national publicatons	Bibliometrics
	8	3	Tools for bibliometics	Bibliometrics
	9	2	Optimal search strategies for databases	IS and IR
	10	2	Information seeking	Bibliometrics
	11	2	Social networking websites	AIT
	12	2	Journal peer review	Bibliometrics
	13	2	National science structures	Bibliometrics
	14	2	Journal impact factors	Bibliometrics
	15	2	Citation analysis	Bibliometrics

Table 5 continued

Year	No. cluster	No. articles	Subject	Broad subject
2010–2014	1	54	Bibliometrics	Bibliometrics
	2	20	E-government and social networks	AIT
	3	17	Coauthorship; h-index	Bibliometrics
	4	13	Scientific collaboration; citation analysis	Bibliometrics
	5	12	Webmetrics; social networking websites	Bibliometrics
	6	10	Literature retrieval system	IS and IR
	7	8	Social networking websites	AIT
	8	7	University ranking	Bibliometrics
	9	4	E-government and information technology	AIT
	10	4	Open access	Other
	11	3	Bibliometric analysis of articles	Bibliometrics
	12	3	Citation analysis	Bibliometrics
	13	2	Recommentation systems	AIT
	14	2	Patentometrics	Bibliometrics
	15	2	Highly cited publications	Bibliometrics
	16	2	Citation bias	Bibliometrics
	17	2	Author productivity	Bibliometrics
	18	2	Authorship	Bibliometrics
	19	2	Research evaluation	Bibliometrics

Table 6 Comparison of broad subjects in the four 5-year periods based on articles with bibliographic coupling relationships

Broad subject	1995–1999		2000–2004		2005–2009		2010–2014	
	No. of clusters	No. of articles (%)	No. of clusters	No. of articles (%)	No. of clusters	No. of articles (%)	No. of clusters	No. of articles (%)
IS and IR	4	41 (66.1)	2	28 (40.0)	3	11 (9.2)	1	10 (5.9)
Bibliometrics	5	21 (33.9)	6	38 (54.3)	9	98 (81.7)	13	121 (71.6)
AIT	0	0	2	4 (5.7)	3	11 (9.2)	4	34 (20.1)
Other	0	0	0	0	0		1	4 (2.4)
Total	9	62 (100.0)	10	70 (100.0)	15	120 (100.0)	19	169 (100.0)

information. In addition, two or more clusters related to citation analysis were included in the bibliometrics broad subject in the four periods, indicating that citation analysis was a critical topic in bibliometrics. When additional topics were added to the bibliometrics research, the category bibliometrics expanded and became an essential subfield within LIS in terms of percentage. Regarding the clusters related to "AIT," the articles focused on search engines, websites, and electronic government.

Table 7 Cluster analysis based on articles with co-citation relationships during the four periods

Year	No. cluster	No. articles	Subject	Broad subject
1995–1999	1	37	Bibliometrics	Bibliometrics
	2	29	Information science, information behavior, and information retrieval	IS and IR
	3	5	Internet information searching by children	IS and IR
	4	4	Relevance judgment	IS and IR
	5	3	Attributes and retrieval of graph files	IS and IR
	6	3	Doctors' information sources	IS and IR
	7	2	Scientometrics	Bibliometrics
	8	2	Performance of search engines	AIT
2000–2004	1	35	Information searching and information retrieval	IS and IR
	2	26	Webmetrics	Bibliometrics
	3	9	Bibliometrics	Bibliometrics
	4	3	E-government	AIT
	5	3	Citation analysis	Bibliometrics
	6	3	Scientific collaboration	Bibliometrics
	7	2	Information needs and seeking in the field of clinical medicine	IS and IR
	8	2	Knowledge management	Other
2005–2009	1	57	H-index	Bibliometrics
	2	17	Visualization of scientific structure; interdisciplinary studies	Bibliometrics
	3	8	Search engines	AIT
	4	6	E-government	AIT
	5	5	Citation analysis	Bibliometrics
	6	3	Bibliometric analysis of Astrophysics Data System	Bibliometrics
	7	2	Information seeking behavior	IS and IR
	8	2	Optimal search strategies for databases	IS and IR
	9	2	Information seeking by team work	IS and IR
	10	2	Social networking websites	AIT
	11	2	Self-citation	Bibliometrics
	12	2	National scientific structure	Bibliometrics
	13	2	Bibliometrics	Bibliometrics
	14	2	Term analysis	Bibliometrics
2010–2014	1	28	Impact factor	Bibliometrics
	2	22	Bibliometric mapping	Bibliometrics
	3	18	E-government	AIT
	4	11	Almetrics; bibliometric indicators	Bibliometrics
	5	8	Scientific collaboration	Bibliometrics
	6	8	University ranking	Bibliometrics
	7	7	Analysis of social networking sites	Bibliometrics
	8	5	Online information communication	IS and IR
	9	5	Retrieval system evaluation	IS and IR

Table 7 continued

Year	No. cluster	No. articles	Subject	Broad subject
	10	5	Scholarly publishing	Bibliometrics
	11	4	Citation analysis	Bibliometrics
	12	3	Authorship	Bibliometrics
	13	3	New bibliometric indicators based on h-index	Bibliometrics
	14	3	Author productivity	Bibliometrics
	15	2	Research evaluation	Bibliometrics
	16	2	Webmetrics	Bibliometrics
	17	2	Patentometrics	Bibliometrics
	18	2	Bibliometric analysis of articles	AIT
	19	2	Open access	Other
	20	2	H-index	Bibliometrics
	21	2	International collaboraton	Bibliometrics

Co-citation analysis

Table 7 shows the clusters generated according to articles with co-citation relationships in the four 5-year periods. In 1995–1999, the two primary clusters were "bibliometrics", containing 37 articles (Cluster 1), and "IS and IR", containing 29 articles (Cluster 2). In 2000–2004, "IS and IR" was the largest cluster, comprising 35 articles. The second largest cluster was "webmetrics," derived from bibliometrics, containing 26 articles. However, a large shift was observed in 2005–2009; the largest cluster, "h-index," contained 57 articles. Although h-index was a new subject within bibliometrics, it was evident that h-index received considerable attention in a short period of time. In 2010–2014, "impact factor" containing 28 articles and "bibliometric mapping" containing 22 articles were the two primary clusters. Among 21 clusters identified, over half of clusters related to bibliometrics. This means that more various bibliometric issues gained substantial visibility than before.

Table 8 shows the distribution of subjects based on articles with co-citation relationships in the four 5-year periods. "IS and IR" and "bibliometrics" were the two main subjects in the first two periods. Although "IS and IR" were not dominant in 2000–2004, the percentage was close to that of "bibliometrics." An increasing trend was identified in the percentage of articles on bibliometrics. Particularly, substantial growth was observed in 2005–2009, during which time "bibliometrics" became the largest broad subject, containing 80.4 % of articles, because of the increasing number of clusters representing more subjects related to bibliometrics. The new cluster, "h-index," played an essential role in enhancing the visibility of bibliometrics. Although the percentage of articles on bibliometrics in 2010–2014 was slightly lower than that in 2005–2009, the percentage was much higher than that of each of other broad subjects. An additional increasing trend was identified in "AIT" with a slight change, whereas a decreasing trend was observed in the percentage of "IS and IR" articles. Because of a sharp decrease in the number of related articles, "IS and IR" became the second smallest broad subject in 2005–2009 and 2010–2014.

Table 8 Comparison of broad subjects in the four 5-year periods based on articles with co-citation relationships

Broad subject	1995–1999		2000–2004		2005–2009		2010–2014	
	No. of clusters	No. of articles (%)	No. of clusters	No. of articles (%)	No. of clusters	No. of articles (%)	No. of clusters	No. of articles (%)
IS and IR	5	44 (51.8)	2	37 (44.6)	3	6 (5.4)	2	10 (6.9)
ibliometrics	2	39 (45.9)	4	41 (49.4)	8	90 (80.4)	16	112 (77.8)
AIT	1	2 (2.4)	1	3 (3.6)	3	16 (14.3)	2	20 (13.9)
Other	0	0	1	2 (2.4)	0	0	1	2 (1.4)
Total	8	85 (100.0)	8	83 (100.0)	14	112 (100.0)	21	144 (100.0)

Comparison of results generated using the three methods

Table 9 shows the results generated using the three methods in the four periods. In 1995–1999, the consistent results generated using the three methods confirmed that "bibliometrics" and "IS and IR" were the two main broad subjects. However, a considerable difference in the percentages of the two broad subjects was identified through bibliographical coupling analysis. This difference was more substantial than the discrepancy in the percentages between the two main broad subjects identified through keyword and co-citation analyses. In 2000–2004, "AIT" was the largest broad subject; however, this was not the result obtained using the other two methods.

In 2005–2009, the main broad subject identified using the three methods was identical: "bibliometrics" was the main subject, with large percentages of 80.4–81.8 %. Regarding "AIT," its percentages were limited, except in the results obtained through keyword analysis in 2000–2004. A reason explaining the limited percentages of "AIT" may be that articles only partially related to the Internet were classified as other broad subjects. The increasing pervasiveness of the Internet has led to a blurring of the boundary between "AIT" and other broad subjects. For example, webmetrics (analysis of data from the Internet) is a subfield of bibliometrics, and "IS and IR" include online searching. In 2010–2014, "bibliometrics" was still the main broad subject according to the results obtained using three methods. However, the percentage of "bibliometrics" was lower than that in 2005–2009.

Discussion

In this study, we used three methods, namely keyword, bibliographical coupling, and co-citation analyses, to explore the evolution of research subjects in LIS in four 5-year periods between 1995 and 2014. The consistent findings generated using the three methods show that three broad subjects, "bibliometrics", "IS and IR", and "AIT", were present in the four periods. Except for in the results obtained using keyword analysis in 2000–2004, "bibliometrics" and "IS and IR" were the two main broad subjects in the first two periods (1995–2004). In addition, a decreasing trend in the percentage of "IS and IR" was confirmed according to the consistent findings generated using the three methods. Substantial changes were identified in 2005–2009. Whereas the prominence of "IS and IR" decreased

Table 9 Comparison of the results regarding the four time periods generated using the three methods

Method	1995–1999		2000–2004		2005–2009		2010–2014	
	Broad subject	%	Broad subject	%	Broad subject	%	Broad subject	%
Keyword analysis	Bibliometrics	35.7	AIT	35.4	Bibliometrics	81.8	Bibliometrics	74.2
	IS and IR	32.8	IS and IR	25.7	IS and IR	8.6	Other	12.3
	Other	16.6	Bibliometrics	24.8	Other	5.2	AIT	10.9
	AIT	15.0	Other	14.2	AIT	4.4	IS and IR	2.6
Bibliographic coupling	IS and IR	66.1	Bibliometrics	54.3	Bibliometrics	81.7	Bibliometrics	71.6
	Bibliometrics	33.9	IS and IR	40.0	AIT	9.2	AIT	20.1
	AIT	0	AIT	5.7	IS and IR	9.2	IS and IR	5.9
	Other	0	Other	0	Other	0	Other	2.4
Co-citation	IS and IR	51.7	Bibliometrics	49.4	Bibliometrics	80.4	Bibliometrics	77.8
	Bibliometrics	45.9	IS and IR	45.9	AIT	14.3	AIT	13.9
	AIT	2.4	AIT	3.6	IS and IR	5.4	IS and IR	6.9
	Other	0	Other	2.4	Other	0	Other	1.4

rapidly in 2005–2009, that of "bibliometrics" increased. The situation extended to the fourth period (2010–2014). The results generated using the three methods revealed that "AIT" exceeded "IS and IR" in number.

Both information seeing and information retrieval are typically interdisciplinary fields. Chang (2011) analyzed 2,003 articles between 1962 and 2009 and reported that studies on information seeking represented 22 disciplines. As many as 66 % of the articles related to information seeking were published in journals of LIS and medical science. Ding et al. (2000) explored the primary fields investigating information retrieval research by using journal co-citation analysis based on journals on information retrieval between 1987 and 1997. They identified four primary fields, namely psychology, LIS and computer science, physics and chemistry, and science. This indicates that information retrieval research is an interdisciplinary field. Articles on information retrieval were not concentrated in LIS journals. As such, the declining trend in the proportion of articles on "IS and IR" may indicate that increasing numbers of "IS and IR" researchers have published articles in fields outside LIS.

The results of both bibliographic coupling analysis and co-citation analysis (but not keyword analysis), demonstrate that "bibliometrics" and "AIT" exhibited increasing trends. Although "bibliometrics" and "AIT" relate to the Internet, the growth rate of "AIT" is distinct from that of "bibliometrics." The pervasive application of the Internet has affected numerous disciplines, causing studies related to "AIT" to appear in various disciplines. Therefore, the low percentage of "AIT" does not indicate that the research on the application of the Internet does not appeal to researchers.

The two studies conducted by Åström (2002, 2007), which elucidated the structure of LIS according to nine journals in 1998–2000 and 21 journals in 1990–2004 respectively, identified "bibliometrics" and "IS and IR" as two subfields within LIS. The obvious difference in the subfields between the present study and those of Åström is that library science was confirmed as another subfield in Åström's studies. The dissimilar findings may be attributable to the various methods used. However, "bibliometrics" and "IS and IR" are evidently two essential research subjects that have existed in LIS and exerted considerable

influence. In particular, bibliometrics has become dominant and considerably influences the research of new topics.

Conclusion

This study identified "bibliometrics" as the most essential research subject in LIS between 1995 and 2014. In addition, "bibliometrics" exhibited increasing influence, although its percentage in 2010–2014 was lower than that in 2005–2009. "IS and IR" was also the primary research subject in the first two periods, however, a decreasing trend was observed, according to trend analysis in four successive 5-year periods. We used articles that were among the top 5 % of highly cited articles. Although the top 5 % of highly cited articles accounted for a limited proportion of all articles, such articles can be regarded as representative of a discipline because of their considerable visibility and influence. Therefore, subject analysis based on highly cited articles involves the most essential subjects. The focus of this study was not to reveal comprehensive subjects explored by LIS researchers.

Because the analyzed units and processes adopted in the three methods differed, we expected differences in the characteristics and trends of LIS research topics among the results of the three methods. Combining various methods benefits researchers who investigate the development of research subjects in specific disciplines. The similarities and differences in the findings obtained using the three methods enable creating a more comprehensive map of the structure of LIS than do the finding obtained using a single method.

The results obtained using bibliographic coupling and co-citation analyses are similar. The advantage of keyword analysis is that it enables presenting specific subjects, revealing subtle subject differences within the same broad subjects. However, certain keywords are so general that they are used in various contexts. In this situation, the subject of an article can be determined with reference to other keywords contained in the article. For example, "evaluation" may refer to system evaluation or research evaluation. The appropriate meaning of evaluation in a specific article depends on other keywords. Regarding bibliographic coupling analysis and co-citation analysis, the analysis unit is a pair of articles. A cluster containing at least two articles is defined based on the subjects of articles. Thus, the scope of subjects generated from bibliographic coupling and co-citation analyses is usually larger than that generated from keyword analysis.

LIS consists of two main subfields, library science and information science, that differ in nature. Most of the journals that were investigated in this study focused on information science; the differences in the nature of information science journals and library science journals may affect the types of research subjects that were identified. This implies that selecting journals representing a specific discipline is an essential process.

References

Åström, F. (2002). Visualizing library and information science concept spaces through keyword and citation based maps and clusters. In H. Bruce, R. Fidel, P. Ingwersen, & P. Vakkari (Eds.), *Emerging frameworks and methods: Processings of the fourth international conference on conceptions of library and information science (CoLIS4)* (pp. 185–197). Westport, CT: Libraries Unlimited.

Åström, F. (2007). Changes in the LIS research front: Time-sliced co-citation analyses of LIS journal articles, 1990–2004. *Journal of the American Society for Information Science and Technology, 58*(7), 947–957.

Bichteler, J., & Eaton, E. A. (1980). The combined use of bibliographic coupling and co-citation for document retrieval. *Journal of the American Society for Information Science, 31*(4), 278–282.

Boyack, K. W., & Klavans, R. (2010). Co-citation analysis, bibliographic coupling, and direct citation: Which citation approach represents the research front most accurately? *Journal of the American Society for Information Science and Technology, 61*(12), 2389–2404.

Chang, H. W. (2004). An analysis of the research areas of articles published in university library journal and college and research libraries between 1998 and 2002. *University Library Journal, 8*(3), 76–83.

Chang, Y. W. (2011). A comparative study of research literature on information needs and information seeking: A bibliometric and social network analysis. *Journal of Educational Media and Library Science, 48*(3), 347–380.

Chen, K. H., & Liang, C. F. (2004). Disciplinary interflow of library and information science in Taiwan. *Journal of Library and Information Studies, 2*(2), 31–55.

Chiu, W. T., & Ho, Y. S. (2007). Bibliometric analysis of tsunami research. *Scientometrics, 73*(1), 3–17.

De-Moya-Anegon, F., Jimenez-Contreras, E., & De La Moneda Corrochano, M. (1998). Research fronts in library and information science in Spain. *Scientometrics, 42*(2), 229–246.

Ding, Y., Chowdhury, G. G., & Foo, S. (2000). Journal as markers of intellectual space: Journal co-citation analysis of information retrieval area, 1987-1997. *Scientometrics, 47*(1), 55–73.

Garfield, E. (1992–1993). New SCI and SSCI CD-ROM editions include abstracts, Keywords plus, and monthly updates. In *Essays of an information scientist* (Vol. 15, pp. 238–247). Philadelphia: ISI Press.

Hu, C. P., Hu, J. M., Gao, Y., & Zhang, Y. K. (2011). A journal co-citation analysis of library and information science in China. *Scientometrics, 86*(3), 657–670.

Jarneving, B. (2005). A comparison of two bibliometric methods for mapping of the research front. *Scientometrics, 65*(2), 245–263.

Järvelin, K., & Vakkari, P. (1993). The evolution of library and information science 1965–1985: A content analysis of journal articles. *Information Processing and Management, 29*(1), 129–144.

Kessler, M. (1963). Bibliographic coupling between scientific papers. *Journal of American Documentation, 14*, 10–25.

Koufogiannakis, D., Slater, L., & Crumley, E. (2004). A content analysis of librarianship research. *Journal of Information Science, 30*(3), 227–239.

Ma, R. (2012). Author bibliographic coupling analysis: A test based on a Chinese academic database. *Journal of Informetrics, 6*, 532–542.

Ma, R., Dai, Q., Ni, C., & Li, X. (2009). An author co-citation analysis of information science in China with Chinese Google scholar search engine, 2004–2006. *Scientometrics, 81*(1), 33–46.

Qiu, J. P., Dong, K., & Yu, H. Q. (2014). Comparative study on structure and correlation among author co-occurrence networks in bibliometrics. *Scientometrics, 101*(2), 1345–1360.

Sharabchiev, J. T. (1989). Cluster analysis of bibliographic references as a scientometrics method. *Scientometrics, 15*(1/2), 127–137.

Shi, M. Y. (2002). *The study of library science in Taiwan based on journal articles*. Taipei: Han-Mei.

Small, H. (1973). Co-citation in the scientific literature: A new measure of the relationship between two documents. *Journal of the American Society for Information Science, 24*(4), 265–269.

Tsai, C. L., Wang, M. H., & Ho, Y. S. (2008). Bibliometric analysis of cataract research from 1991 through 2006, analyzed with the science citation index. *Acta Societatis Ophthalmologicae Sinicae, 47*(2), 101–111.

Tsay, M. Y., & Hsu, Y. T. (2009). An analysis of subjects and citations in core Journals of library and information science in Taiwan and China from 1997 to 2006. *Journal of Librarianship and Information Studies, 1*(3), 17–37.

Tsay, M. Y., & Lai, C. H. (2007). Comparison of literature scattering and subject changes between citing and cited literature on information science: 1985–2005. *Journal of Library and Information Science Research, 1*(2), 1–32.

White, H. D., & Griffith, B. C. (1981). Author co-citation: A literature measure of intellectual structure. *Journal of the American Society for Information Science, 32*(3), 163–171.

White, H. D., & McCain, K. W. (1998). Visualizing a discipline: An author co-citation analysis of information science, 1972–1995. *Journal of the American Society for Information Science, 49*(4), 327–355.

Yan, E., & Ding, Y. (2012). Scholarly network similaries: How bibliometric coupling networks, citation networks, citation networks, co-citation networks, topical networks, co-authorship networks, and co-word networks relate to each other. *Journal of the American Society for Information Science and Technology, 63*(7), 1313–1326.

Yi, H., Ao, X., & Ho, Y. (2008). Use of citation per publication as an indicator to evaluate pentachlorophenol research. *Scientometrics, 75*(1), 67–80.

Zhao, D., & Strotmann, A. (2008). Evolution of research activities and intellectual influences in information science 1996–2005: Introducing author bibliographic coupling analysis. *Journal of the American Society for Information Science and Technology, 59*(13), 2070–2086.

Zhao, D., & Strotmann, A. (2014). The knowledge base and research front of information science 2006–2010: An author co-citation and bibliographic coupling analysis. *Journal of the Association for Information Science and Technology, 65*(5), 995–1006.