Changes in the LIS Research Front: Time-Sliced Cocitation Analyses of LIS Journal Articles, 1990–2004

Åström, Fredrik

Published in:
Journal of the American Society for Information Science and Technology

DOI:
10.1002/asi.20567

2007

Changes in the LIS Research Front: Time-Sliced Cocitation Analyses of LIS Journal Articles, 1990–2004

Fredrik Åström
Department of Cultural Sciences, BIVIL, Biskopsgatan 7, SE-223 62 Lund, Sweden.
E-mail: fredrik.astrom@kult.lu.se

Based on articles published in 1990–2004 in 21 library and information science (LIS) journals, a set of cocitation analyses was performed to study changes in research fronts over the last 15 years, where LIS is at now, and to discuss where it is heading. To study research fronts, here defined as current and influential cocited articles, a citations among documents methodology was applied; and to study changes, the analyses were time-sliced into three 5-year periods. The results show a stable structure of two distinct research fields: informetrics and information seeking and retrieval (ISR). However, experimental retrieval research and user oriented research have merged into one ISR field; and IR and informetrics also show signs of coming closer together, sharing research interests and methodologies, making informetrics research more visible in mainstream LIS research. Furthermore, the focus on the Internet, both in ISR research and in informetrics—where webometrics quickly has become a dominating research area—is an important change. The future is discussed in terms of LIS dependency on technology, how integration of research areas as well as technical systems can be expected to continue to characterize LIS research, and how webometrics will continue to develop and find applications.

Introduction

Since the 1970s, the role and conceptualization—as well as the management of—information have gone through substantial changes, a development intensified by the expansion of the Internet and the World Wide Web (WWW) in the 1990s. One effect of this has been an increased interest in information-related issues in several fields of research such as management studies, as well as the computer sciences, for example. For library and information science (LIS), this has meant increased competition; the question is how LIS has responded to this. What has been on the frontline of LIS research over the last 15 years?

Received January 11, 2006; revised July 11, 2006; accepted July 11, 2006

 • Published online 12 March 2007 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/asi.20567

The nature and intellectual organization of LIS has been thoroughly investigated in analyses describing the general traits of LIS research, as well as mapping how LIS has been organized in different research themes (Persson, 1994; White & Griffith, 1981; White & McCain, 1998). However, fewer studies have dealt with the period after 1995—one exception is Morris (2004) visualization of LIS 1981–2001—and few have analyzed changes within the discipline over time. White and McCain (1998) study three 8-year periods, however, primarily oriented towards author movements between research areas. Furthermore, Persson's (1994) analysis is based on articles from one journal, whereas White and McCain (1998) cover 12 journals with a strong dominance of Information Science journals. A wider journal selection, especially one including more journals with a Library Science orientation, has been argued for by Åström (2002).

Informetric analyses on research fronts are well established (e.g., Garfield, 1994; Morris, Yen, Wu, & Asnake, 2003; Persson, 1994; Price, 1965); in addition, there have been attempts at predicting the ranking of papers and authors by analyzing research fronts (e.g., Burrell, 2003; Feitelson & Yovel, 2003; Glänzel, 1997; Glänzel & Schoepflin, 1994; Glänzel & Schubert, 1995). Furthermore, there are also examples of studies of research fields where analyses of trends have been used to predict future development of research fields (e.g., Yashin & Yashin, 2001). In LIS, these issues has also been discussed by Persson (2005), analyzing the informetrics field through articles in *Scientometrics*.

In this article, my aim is to analyze changes in LIS research over the last 15 years. My approach centers on the following questions. What research topics have dominated LIS during the period 1990–2004? What changes can be observed in the topics addressed over the last 15 years? Can these changes can be used to tell us something about where LIS is heading?

Most definitions of "research fronts" explain them as groups of citing articles being clustered through bibliographic coupling (e.g., Persson, 1994), and their relations to the cited documents clustered by cocitation analysis (Garfield, 1994; Morris et al., 2003; Price, 1965). Although

Persson sees the current (citing) articles as the research front and the cited documents as the research base, Garfield, for example, also includes the clusters of cocited core articles into the research front. The distinction between research base and front suggested by Persson (1994) is maintained in this article; the research base clusters are determined by an unrestricted cocitation analysis. What constitutes the research front, however, is determined by a restricted cocitation analysis. By limiting the analysis to cocitations within a temporally limited document set, the currentness strived for is maintained. In addition, by analyzing the co-occurrence of highly cited documents, we also get an indication on the impact of the articles, thus expanding the definition of research fronts as including influential, as well as current research.

Method

To identify LIS research, and to select journals for the analyses, the *Journal Citation Reports: JCR Social Sciences* (Thomson ISI, 2003) was used. To defining LIS research, *JCR*'s Information Science & Library Science classification, covering 55 journals, was used. To limit the definition, all general LIS journals were identified and the specialized ones were excluded. This was done using the "Citing Journal" field in *JCR*: If the journal primarily was cited by non-LIS publications, it was excluded from the study. This is motivated by a need to limit the analysis to LIS specifically, i.e., to journals within the specialty and journals in the directly adjacent non-LIS field. The specialty can be seen as part of a much larger field, but its interaction with the larger LIS field is limited and therefore beyond the scope of this investigation.

The focus of the study is on research articles per se; hence, the next step was to exclude the yearly review series *Annual Review of Information Science and Technology,* and a few trade journals such as *Online.* By selecting all general LIS journals in the *JCR*, the exclusion of Library Science-oriented journals was avoided. In the final step, journals published throughout the whole period 1990–2004 were selected, providing 21 journals for inclusion in the study (Table 1). From these journals, the ISI Web of Science was used to download all 13,605 genuine research articles 1990–2004.

The data downloaded from Web of Science was processed and analyzed using the Bibexcel software (Persson, 2006). In all analyses, the highest cited documents were selected; and in-between those, the raw cocitation frequencies were computed. The matrix with all the cocitation frequencies was processed using the multidimensional scaling algorithm, MDS Alscal. The result is a two-dimensional map, where the circles represent the cited articles analyzed; the distances between the circles are based on the raw cocitation frequencies. How accurate the MDS computes the distances is indicated by the stress value, where lower values indicate a better fit between the map and the cocitation matrix. The lines between the articles represent the strongest cocitation links, included to enhance the structures of the map. These links were produced using a clustering routine suggested by Persson (1994). Apart from identifying the cocited pairs, the clustering routine also requires at least one unit in another cocited pair being linked to either of the articles in the first pair to form a cluster. What cluster the individual documents belong to has been marked by adding the cluster number to the document labels.

The analyses were done in three steps. The first step was a cocitation analysis done on the whole set of articles from 1990–2004, to determine the research base; this was done without any limitations on the cited side. The second step was another cocitation analysis, again on the whole document set, but this time the cited side was limited by doing a citations among documents (CAD) analysis. This was done to identify how LIS research has been structured during the last 15 years, thus providing a research context for the whole 15-year period. The CAD analysis limits the number of cited documents to those occurring within the same document set as the citing documents, using a search key to identify citations to documents within the set (Persson, 2001). In addition, self-citations were removed to concentrate the analysis to connections between documents as seen by other authors. The third step incorporated three cocitation analyses, again using the CAD analysis to limit the time window on the cited side. These three analyses were also time-sliced, to identify the research fronts and discover changes in and between research areas. Time slicing is performed by extracting documents published during a certain period within the document set, thus creating a subset to analyze (Chen, 2004; Persson, 2005). Here, three subsets were created, covering the periods 1990–1994, 1995–1999, and 2000–2004.

The analyses were done on a document level, as opposed to an analysis on the author level. Although an author analysis provides more of an overview, the document analysis is

TABLE 1. Journals selected for download and analysis.

Aslib Proceedings	*Journal of Documentation*	*Library Resources & Technical Services*
College & Research Libraries	*Journal of Information Science*	*Library Trends*
Electronic Library	*Journal of Librarianship/. . . and Information Science*	*Libri*
Information Processing & Management	*Journal of the American Society for Information Science/. . . and Technology*	*Proceedings of the ASIS/ASIST Annual Meeting*
Information Technology and Libraries	*Library & Information Science Research*	*Scientometrics*
Interlending & Document Supply	*Library Journal*	
Journal of Academic Librarianship	*Library Quarterly*	

DOI: 10.1002/asi

more detailed, e.g., by not grouping documents on different topics by the same author. The main reason for selecting the CAD analysis is to keep a high level of concentration, temporally and in terms of topic matter. An unrestricted cocitation analysis was done without limiting the cited documents included in terms of when or where it was published, or from which field or discipline the research originated. Thus, the research base used in this analysis includes all documents that form the intellectual foundation of the citing set of documents, regardless of age, scholarly origin, or form of the document. To study research fronts, however, the analysis needs to focus on contemporary research, while avoiding topic drift. This is achieved by the use of a restricting CAD methodology, limiting the set of *cited* documents to only include those also among the *citing* documents. Thus, topic drift is avoided by only studying cocitations among documents from the same—topically related—journals. A temporal restriction is also enforced because the analysis includes only cited documents contemporary to the citing documents. The result reflects contemporary and influential research within a specific field of research, i.e., the research front.

Another advantage is that CAD analyses also make it possible to control for self-citations (Persson, 2001), that is, "whenever the set of [authors or] co-authors of the citing paper and that of the cited one are not disjoint, that is, if these sets share at least one author" (Glänzel, Thijs, & Schlemmer, 2004, p. 65). Self-citations are debated within the bibliometric literature: on one hand they can be used for inflating citation counts and creating a stronger position in the scientific community, on the other hand some self-citations can be seen as a natural part of the scientific communication process (Glänzel, Thijs, & Schlemmer, 2004). In the analyses presented here, self-citations were removed because research fronts as such are influential research areas within a certain period, and the potential inflation of self-citations might bias research areas. Preliminary analyses were done without removing the self-citations, showing no variations in how the research areas were organized. However, some documents were ranked substantially higher when self-citations were allowed; however, this was not important in defining the clusters.

The reason for performing the time-sliced CAD analyses is to discover changes in and between the different LIS research areas over time. The results show the most cited documents within the subsets, and their relations to each other, telling us what research areas has been growing or declining during the fifteen years; and thus reflecting changes in the research fronts in LIS during the three periods analyzed.

Results

Two analyses covering the whole period were done to give a historical overview of the research base for LIS research from 1990–2004, and to place the research base in a wider context by identifying influential research areas and documents over the years 1990–2004 for comparative purposes.

Research Base

The research base was based on the 13,605 journal articles published from 1990–2004 and their 221,586 references to 150,145 unique documents. The 66 most-cited documents that received 50 citations or more were selected for further analysis (Figure 1).

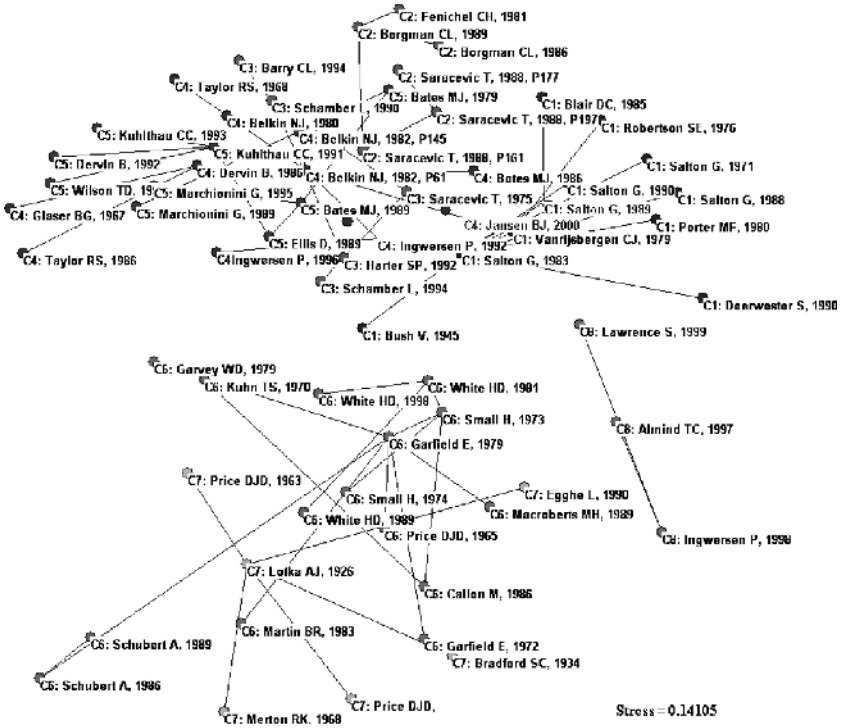

FIG. 1. Cocitation map of library and information science (LIS) research base showing articles cited in LIS journals, 1990–2004.

DOI: 10.1002/asi

The map shows two main areas consistent with the structures found in earlier analyses on LIS (e.g., Persson, 1994; White & McCain, 1998). On the top half of the map, a group of information-seeking and retrieval (ISR) related literature is featured and on the bottom half, a group of informetrics literature. However, on the right side of the informetrics field, a group of webometric studies has formed a cluster. Webometrics is the study of the nature and properties of the World Wide Web, using informetric methodologies such as link, citation, and cluster analyses (Björneborn & Ingwersen, 2001). Although literature on the WWW and webometrics is fairly new, it has already managed to take its place in the research base in LIS.

In the ISR section of the map, there is a thematic shift from right to left. Systems-oriented information retrieval (IR) literature is on the far right, followed towards the left by user-system interaction studies and information behavior. In comparison to Persson (1994), the "soft" part of the IR-field has increased its impact compared to the "hard" systems-oriented IR research. This shift also has a temporal dimension: time lining the evolution of ISR research. Documents on the right side are primarily from the 1970s and 1980s, the middle part from the 1980s and early 1990s, and the left side is dominated by literature from the 1990s. Apart from the webometric group on the far right, the informetrics field is centered on bibliometric mapping, surrounded by documents concerning bibliometric distributions.

To enhance the results of the cocitation analysis, a cluster analysis (Persson, 1994) was performed, resulting in eight clusters (Table 2). The clusters support the structures identified in the map, and reveal a division of the soft IR-research: from search- and relevance-focused documents, over cognitive IR and information seeking, to information behavior. There are several overlaps in the clusters, but it reflects the development of research trends that has been visible in LIS research over the last 30–40 years.

There are major differences between the clusters, both in terms of how many citations they have attracted and how many documents are included. Of the cluster with documents placing themselves on the top half of the map, the experimental IR cluster is still by far the strongest; it is also the dominating cluster when looking at citation frequencies. However, when adding the information searching and seeking clusters, they receive almost twice as many citations as the systems-oriented cluster.

The publication years of the clustered documents shows four generations of research orientations, a trait also visible in the IR part of the map. The first generation of LIS research includes experimental IR, bibliometric distributions, and bibliometric mapping. The second generation of research, with references published from the early 1980s marks the increasing interest in the user side of IR, incorporating the search process and the cognitive perspective into IR and LIS research. This is followed by the relevance studies in the early 1990s; and a contemporary trend to focus on general information behavior. The most recent trend in the LIS research base is studies on World Wide Web and webometrics, dating back to the late 1990s.

The LIS research base map consists of well-defined clusters; and the structures and content are also very stable in comparison to previous studies, with the addition of webometrics. The stability is also supported by the low frequency of documents from the mid-1990s or later, revealing a consistency in the literature current LIS research is based on.

Research Context 1990–2004

The results of the second analysis show influential research areas during the period 1990–2004. It is still the same 13,605 articles providing the material, but only the 18,615 citations to articles present in the set of citing documents are analyzed. Here, as well as in the following time-sliced analysis, the self-citations were removed. Out of the 5024 unique-cited documents, the 65 articles being cited 25 times or more were selected and analyzed (Figure 2).

The general structure of the map is the same: with informetrics on the lower half and ISR on the top half. There are some differences, however. In the top half, a center has developed around "Kuhlthau, 1991" and "Ingwersen, 1996," focusing on cognitive ISR, relevance, and information behavior, while experimental IR research has become peripheral. Different perspectives on the user-oriented research has dominated the

TABLE 2. Document distribution, citation frequency, and median age in library and information science (LIS) research base clusters cited in LIS journals, 1990–2004.

Cluster	No. of documents	No. of citations	Median publication year
Cluster 1. Experimental information retrieval (IR)	11	1239	1983
Cluster 2. IR/Information search	6	409	1988
Cluster 3. IR/Relevance	5	397	1992
Cluster 4. Information seeking and use/Cognitive IR	11	855	1986
Cluster 5. Information seeking and use/information behavior	9	692	1991
Cluster 6. Bibliometric mapping	15	1036	1981
Cluster 7. Bibliometric distributions	6	455	1968
Cluster 8. World Wide Web/webometrics	3	188	1998
Total	66	5271	1986

DOI: 10.1002/asi

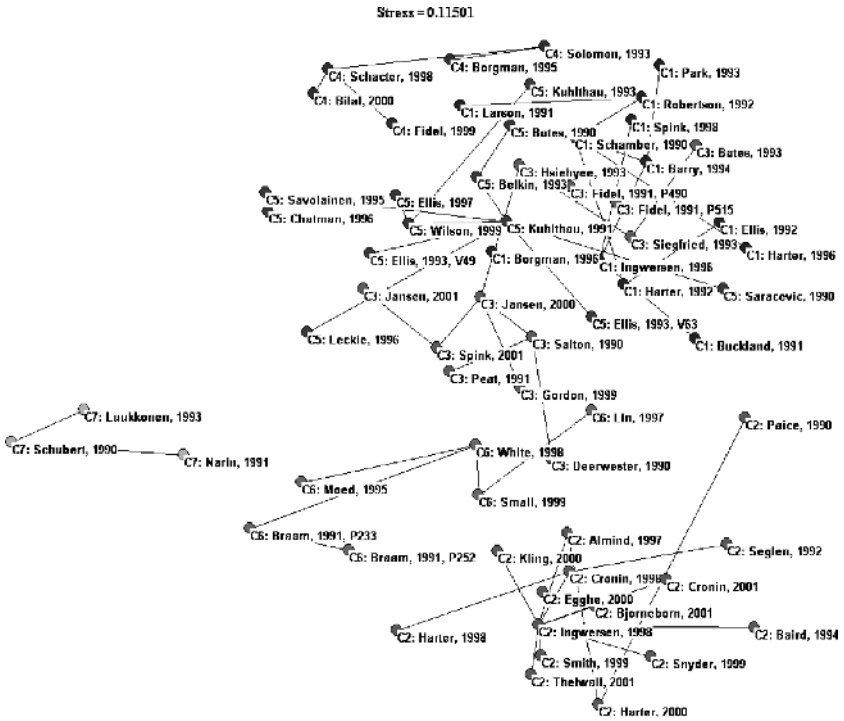

FIG. 2. Cocitation map of library and information science (LIS) research, 1990–2004, based on a LIS citations among document analysis of LIS journals.

information-seeking and retrieval field; and has together with the wider information behavior field formed a strong research area of different variations on information-seeking research.

At the same time, the informetrics field has become more dispersed, with three clearly defined subfields: research collaboration to the left, bibliometric mapping in the middle, and webometrics on the right side. In comparison with the research base, webometrics has become the dominating research area within the informetrics field.

The difference between the research base and context clusters is: in the research context analysis, the experimental IR and the bibliometric distributions clusters disappeared, while two new clusters entered: studies on children's information behavior and informetric analyses on research collaboration (Table 3). There are still documents concerning hard IR and bibliometric distributions, but not so many that they form clusters on their own. Instead, they are incorporated into other clusters such as the IR/Search cluster.

Another aspect is the distribution of citations in-between the clusters, where they are more evenly spread between clusters in the research context than in the research base. Whereas the research base has two heavily dominating clusters, the exceptions in the context analysis are three smaller clusters: two about half the size and one about 20–25% of the size of the majority of the clusters.

The informetrics field has three dominating foci: mapping, research collaboration, and webometrics. The clusters are less unstable, and the correspondence between clusters and titles matches more than in the ISR group. Compared to the research base, there are not as many clusters that are limited to a certain period.

TABLE 3. Document distribution and citation frequency in library and information science (LIS) research clusters in, and cited by LIS journals, 1990–2004.

Cluster	No. of docs.	No. of citations
Cluster 1. Information retrieval (IR)/Relevance	12	555
Cluster 2. World Wide Web/ Webometrics	15	543
Cluster 3. IR/Information search	12	497
Cluster 4. Children's information behavior	5	174
Cluster 5. Information seeking and use/information behavior	12	480
Cluster 6. Bibliometric mapping	6	208
Cluster 7. Informetrics/research collaboration	3	92
Total	65	2549

Research Front 1990–1994

Within the whole document set, 3401 articles are from the years 1990–1994, containing 57,843 references. Of these, 1667 citation links are between articles within the 1990–1994 subset. After removing self-citations, 1581 citations were made to 852 references. The 39 articles receiving five citations or more were selected for analysis (Figure 3).

In the first time-sliced analysis, the basic structure still remains with an informetrics field and an ISR field, although the ISR area is more dominating. The ISR part is more collected than in the 1990–2004 research context analysis. However, there is one small group at the left top part standing

DOI: 10.1002/asi

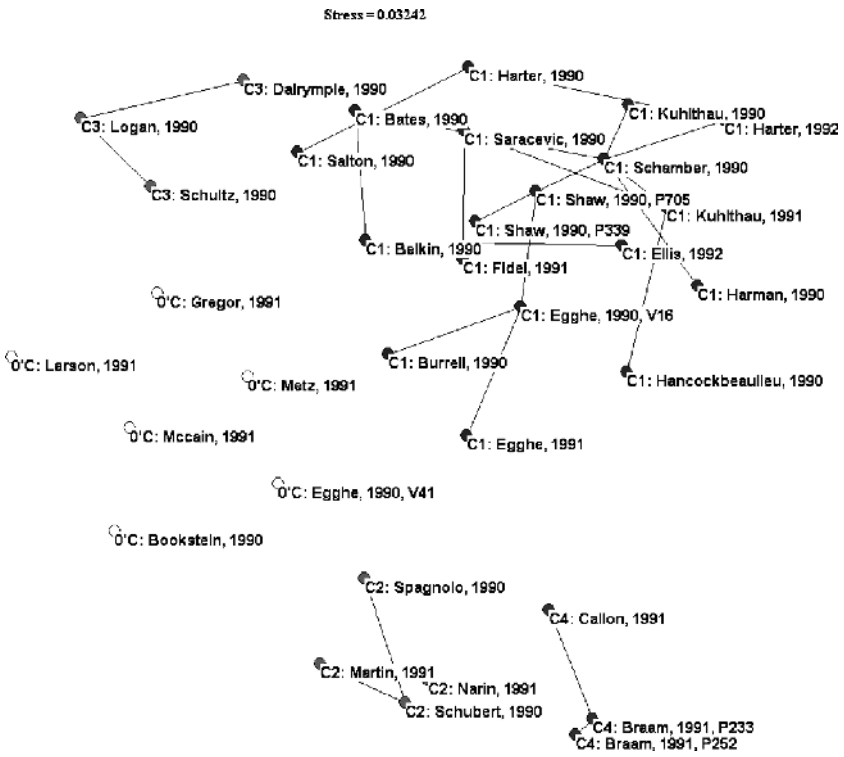

FIG. 3. Cocitation map of the library and information science (LIS) research, 1990–1994, based on a time-sliced citations among document analysis of LIS journals.

TABLE 4. Frequency and document distribution in library and information science (LIS) research front clusters, 1990–1994.

Cluster	No. of documents	No. of citations
Cluster 1. Information seeking and retrieval	18	166
Cluster 2. Informetrics/ Research collaboration	4	31
Cluster 3. Information retrieval/ information search	3	18
Cluster 4. Bibliometric mapping	3	18
0 Cluster	6	35
Total	34	268

out; while the informetrics field is clearly divided into two distinct areas of mapping and research collaboration.

Most of the ISR documents are gathered in one big cluster; a cluster that gathers more than half of the documents analyzed and also more than 60% of the distribution of citations (Table 4). With the exception of a small information search-oriented cluster, the big ISR cluster gathers all variations of IR and information-seeking research. As the map also shows, the informetrics field is divided into two clusters, one oriented towards research collaboration and one focusing on mapping.

There are six documents not making it into any clusters, failing to meet the criteria for inclusion in the clustering routine. The inclusiveness of the IR clusters might leave some doubts on the effectiveness of the clustering routine, but it also reflects developments within the field from the early 1990s and onwards, pointing toward an integration of systems and user-oriented research.

The research front in the years 1990–1994 is heavily dominated by information seeking and retrieval issues, especially from a user perspective. More than half of the documents included in the analysis have a user perspective on information seeking and retrieval; and almost half of them deal in one way or the other with user-system interaction. About 25% of the documents also cover behavioral issues in information-seeking and retrieval processes. Out of all the documents included in the analysis, only about 10% have a systems perspective. It should also be noted that different categorizations of the ISR research are overlapping, whereas in the substantially smaller informetrics field, the divisions are relatively clear-cut. There are no overlaps between the groups; and the number of documents and citations are also quite equally distributed. There are, however, some connections between some informetrics issues and some IR issues, reflecting similarities in terms of research material and methods.

Research Front 1995–1999

The document set contains 3318 documents published from 1995–1999, citing 68,438 documents. After removing self-citations, 2117 citation links were found, referring to 1013 unique documents within the 1995–1999 subset. Of these, the 52 most cited (more than five citations) articles were selected and analyzed (Figure 4).

As in the 1990–1994 analysis, the ISR group is well gathered, with two exceptions on the left side of the map: one focusing on information technology and one on hard IR. The

DOI: 10.1002/asi

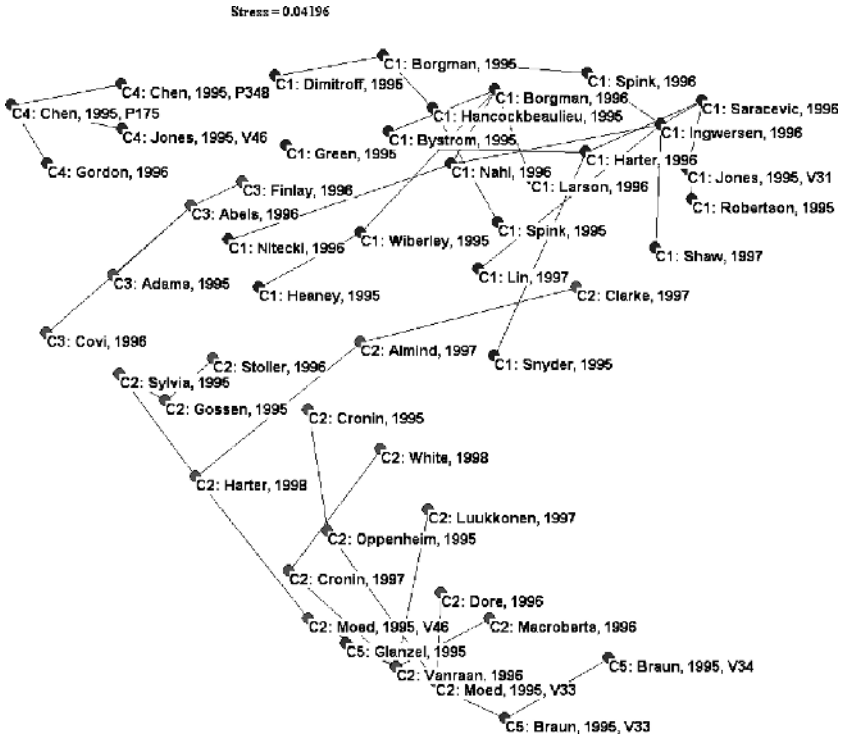

FIG. 4. Cocitation map of the library and information science (LIS) research, 1995–1999, based on a time-sliced citations among document analysis of LIS journals.

TABLE 5. Frequency and document distribution in library and information science (LIS) research front clusters, 1995–1999.

Cluster	No. of documents	No. of citations
Cluster 1. Information seeking and retrieval	21	201
Cluster 2. Informetrics	16	123
Cluster 3. Information Technology	4	34
Cluster 4. Experimental information retrieval	4	28
Cluster 5. Bibliometrics/research performance	3	27
Total	48	413

informetrics field is becoming less dispersed, no longer divided into definable groups as in the earlier analyses. The integration of both the ISR and informetrics field is apparent in the clusters, where large general clusters have developed (Table 5).

Apart from the two general clusters gathering ISR and informetrics research, there are three more-specialized clusters: a research performance cluster in the informetrics field, an experimental IR cluster, and a new cluster with information technology research. These clusters are very small, but there are also documents in the general clusters with titles suggesting they should be part of the specialized clusters.

Again, the ISR field is dominating the 1995–1999 research front, although not as much as in 1990–1994. The broad-spectrum nature of the main clusters suggests little about discrepancies between the clusters and the titles included, although the titles suggest the possibility of more-specialized clusters. A closer look at the documents included in the analysis reveals some interesting issues. As in the 1990–1994 analysis, the user perspective is dominating the ISR field, but at the same time, systems-related issues seem to have gained momentum. There are also fewer documents dealing with behavioral issues, but at the same time, these documents are highly cited.

Another issue relating to the wider ISR field is how the documents gathered in the hard IR cluster all have strong connections to the knowledge organization (KO) field, increasing the visibility of KO research. Whereas earlier documents relating to systems-oriented IR was focused on algorithms for ranking search results, the documents in this period are more focused on organizing the documents and building thesauri.

The emergence of an information technology (IT) cluster is a consequence of LIS interest in IT throughout the 1990s, especially during the latter half, following the breakthrough of the World Wide Web in 1995.

Research Front 2000–2004

The 2000–2004 subset has 4147 records containing 78,646 citations. Of those, 2926 are to documents within the subset. After removing self-citations, 1044 unique documents remain, receiving 2702 citations. The 62 documents receiving seven citations or more were selected for analysis (Figure 5).

DOI: 10.1002/asi

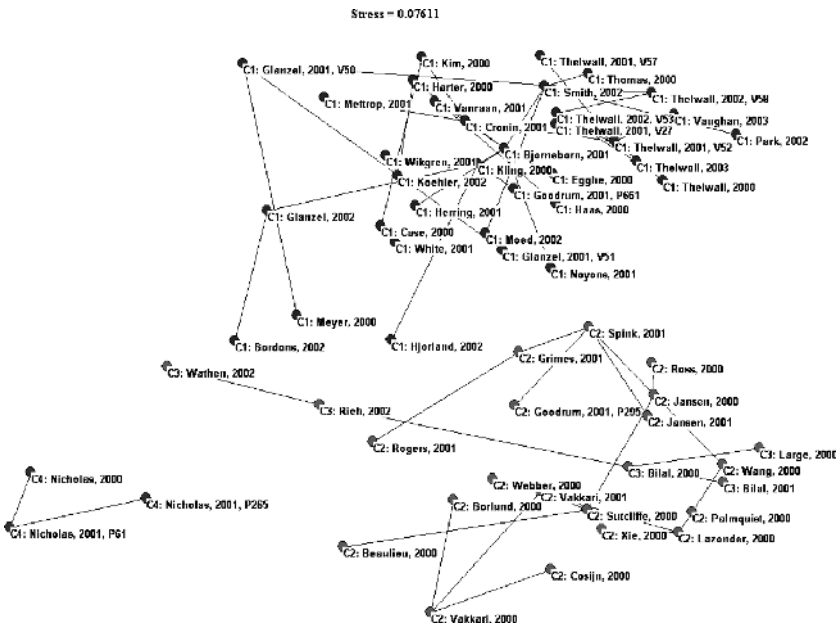

FIG. 5. Cocitation map of the library and information science (LIS) research, 2000–2004, based on a time-sliced citations among document analysis of LIS journals.

TABLE 6. Frequency and document distribution in library and information science (LIS) research front clusters, 2000–2004.

Cluster	No. of documents	No. of citations
Cluster 1. Webometrics	33	472
Cluster 2. Information seeking and retrieval	19	275
Cluster 3. Children's information behavior	5	65
Cluster 4. Health informatics	3	31
Total	60	843

In the top half of this map, the informetrics field is even more closely connected and centered. Moreover, although the field was rather diverse in 1995–1999, in 2000–2004 it is predominantly focused on webometrics. The ISR field on the other hand, has become more separated; at the bottom left of the figure, there is also an isolated health informatics group. The clusters also support this by gathering all informetrics literature in one cluster; while the ISR field is divided into three groups (Table 6).

There are two new clusters emerging in the 2000–2004 analysis: information behavior among children, which was also present in the 1990–2004 context analysis, and health informatics, a small and isolated cluster. In addition, for the first time the informetrics cluster is both larger and more heavily cited than the ISR cluster/clusters. The informetrics field has turned into one big webometrics cluster. About two thirds of the documents included in the cluster are oriented towards webometrics, receiving almost 75% of the citations; and the cluster formed also includes all other informetrics documents. The cocitation frequencies and the distances provided by the MDS Alscal algorithm reveals some distance between the webometrics literature and documents oriented towards research evaluation. However, when applying the clustering routine, the cocitation frequency and amount of "traditional" informetrics documents is not high enough to form a cluster on its own. Within the cluster, the focus is on webometric modeling, although studies on research performance are also an important part of webometrics research.

When looking at the titles of the documents, it also becomes apparent that the impact of the Internet and the WWW is not limited to the webometrics cluster. A majority of the documents in the ISR cluster—and all documents in the children's information behavior cluster—are also focused on the Internet. Another aspect of the ISR field is how behavioral issues again are being cited more frequently. At the same time, most of these are related to information search and not to general information-seeking processes.

Discussion

The analyses presented here, as well as previous studies, show a stable LIS structure over time. The main fields are information seeking and retrieval (ISR) and informetrics, and changes in the discipline can be seen primarily within these fields, not in new fields entering the discipline. There is one substantial change compared to previous analyses: what used to be an IR field has now turned into an ISR field.

Another general observation is how the research base to a much larger extent is dominated by theoretical and methodological texts, whereas the research context and fronts primarily consist of documents of a more empirical nature. The exception is the webometrics area, where webometric

DOI: 10.1002/asi

modeling is dominating over applications or empirically oriented foci. The research base is by definition of a more theoretical and methodological nature, providing the background for empirical and applications-oriented studies. Therefore, most documents of a general theoretical or methodological nature in the time-sliced analyses are found in new research areas.

One of the most striking aspects is the total domination of the WWW as research object in 2000–2004. In 1990–1994, libraries and information services were the main forum for information search; and 1995–1999 was dominated by online databases and information seeking as a broader concept. About two thirds of the documents from 2000–2004 deal in some way with the WWW, from studies on how groups use the Web for searching for information to developing methods for analyzing impact factors for Web sites.

The research base for the ISR field has three basic themes: experimental IR focusing on the development of algorithms, one user-oriented theme focusing on user-system interaction, and another focusing on the user and its behavior in a broader information-seeking process. Compared to the research base, the time-sliced analyses reveal how experimental IR almost disappears as an individual research area, also supported by the lack of hard IR documents published after 1990 in the research base. On the cluster level, experimental IR has merged, primarily with the user-system interaction research. This reflects the general development in the ISR field, where all three ISR areas found in the research base have started merging into one big research area. The idea to use knowledge about user behavior and seeking processes is not new; but during the 1990s, the technological infrastructure has made the development of more flexible systems possible. This is also reflected in the development of information behavior studies during the 1990s. In the early 1990s, this area was strong, and largely theoretically oriented. In the mid-1990s, the area saw a relative decline followed by a reemergence in the late 1990s and the 2000s, this time though, more empirically oriented and relating to systems and systems development. The notion of an integrated ISR area is also supported by Ingwersen and Järvelin (2005) in their discussions on the turn in LIS research, where the information-seeking and IR areas are integrating to create better conditions for the dissemination of information.

The emergence of a health informatics area is interesting; previous studies have shown no connection between LIS and medical informatics (Morris & McCain, 1998). The inclusion of the area can be explained by the analysis being performed on a document level; the area includes three documents by the same author and would disappear if the analysis was made on the author level or—in the case of Morris and McCain (1998)—the journal level.

The informetrics field is dominated by three main research areas in the research base and during the 1990s, a strong research area focused on research performance and collaboration; all four areas merging over time. In 1990–1994, mapping and research performance are distinct research areas, while in 1995–1999, informetrics research has merged into one general area with several research foci combined with one independent area focusing on research performance. In 2000–2004, the whole informetrics field has become one large webometrics area. There are still about one third of the documents dealing with other orientations, but the strength of the highly cited webometrics documents pull all other informetrics texts into the cluster.

One final observation is how the ISR and informetrics fields have closed in on each other. When comparing the research base and the 1990–1994 analysis with the 1990–2004 research context analysis and the two latter periods, the distance between the fields has decreased substantially, and there are overlaps in the links. The empirical material is largely the same, and they both share basic assumptions about, e.g., how to measure the strength of relations between documents or search questions. From the 1990s and onwards, there has also been a development of mutual interests such as visualization of information and an increased migration of research methods between the two fields. This also indicates a stronger visibility of informetrics in LIS mainstream research.

Conclusion

One thing making it hard predicting the future of LIS is that LIS by its nature is technology-driven (Saracevic, 1999). Many research questions are born out of available technology: from the applications of computers for bibliographic purposes giving birth to IR research via the impact of the *Science Citation Index* on informetrics research, to the recent focus on Web-oriented research and the impact of flexible systems on the ISR integration. For now, Web-oriented research can be expected to increase further. One interesting question is whether mobile information technology will have any noteworthy influence on LIS.

One research area that has grown considerably, to the extent that it has come to dominate LIS research over the last 5 years, is webometrics. So far, the main interest has been directed in studies of Web impact factors and Web link analyses, research trying to identify the basic components and structures of Web links and Web sites as material for informetric analyses. When looking at the research tradition of informetrics research in general, this can be compared to the bibliometric distributions area in the research base, laying the foundation for different applications of informetrics research in terms of, e.g., mapping and studies of research performance and so on. There are already examples of these kinds of studies, and the area can be expected to grow further.

The impact of the Web as a field of interest in LIS research is also evident in the information-seeking and retrieval field, where studies on Internet use are as dominating in the ISR field as is webometrics in the informetrics field. This is hardly surprising, considering the impact of the Web on the acquisition of information both in the personal and professional life of people. The Internet is now the main channel for information searching in private life; and most

DOI: 10.1002/asi

professional channels of information such as databases and library catalogues, as well as information infrastructures within organizations, are nowadays available through Web interfaces or in Web-based formats.

The merging of many different channels of information into one medium serves as one of many examples of different kinds of integration that has been evident in LIS research over the last 10–15 years, making merging and integration important concepts in LIS research and services. One aspect is the aforementioned integration of the different research areas within the ISR field, where user-oriented research has started to have an increased impact on the systems-oriented IR research and vice versa. This development is likely to continue. One reason is the continuing development of, and increasing accessibility to, flexible systems capable of accommodating different users. Another reason is recent attempts at formulating a basis for research on the integration of the ISR area (Ingwersen & Järvelin, 2005). Another aspect of integration is the merging of ISR and informetrics research, where an increased interest in methodologies and mapping techniques developed in informetrics have had an impact on IR systems research, something that will continue to have a big impact on retrieval studies in general, but in WWW situations in particular.

The maps show structures and groupings of documents that make sense when compared to LIS research over the last 15 years. As mentioned before, the general structure is stable: both in relation to previous analyses and when comparing the research base with the times-sliced analyses. However, the analyses also show important changes within these fields, e.g., the gradual integration of the ISR field and the emergence of webometrics.

This is also reflected in the clusters, although in a more problematic way. The main purpose of the clustering routine is to support the interpretation of the maps; it has also been a useful tool for defining research areas within the two main fields. However, the clustering routine is sensitive to the size of the data being analyzed, and to the distribution of citations within the document set. When dealing with the whole 1990–2004 period, the clustering is very accurate, whereas the 1990–1994 analysis is less precise. A related issue is how highly cited documents pull in other documents without any strong intellectual connection, something that can be exemplified with the webometrics cluster in the 2000–2004 analysis. A majority of the documents in the cluster is oriented towards webometrics, but there is also a group of other informetric analyses, as well as a few with very different orientations. On the other hand, this can also be used to emphasize the strength of the documents dominating the different clusters, but the pull might also be possible to avoid by setting a higher citation threshold. However, the inclusion of documents that might seem out of place can also, to some extent, be explained by the heterogeneous nature of LIS.

The time-sliced analyses shows some interesting results when analyzing changes in the LIS research fronts over the last 15 years, like the merging of the ISR field, the IR and informetrics fields coming closer to each other; and of course, the development of Web-oriented studies both in ISR and informetrics research. There are some possible problems with using discrete timeframes though. With the fixed 5-year periods, there is the risk of missing important information because documents published at the end of the period do not have time to gather enough citations to become part of the time-sliced analysis, even though they might rank high in the analysis of the total document set.

The exclusion of self-citations by use of the citation among document analyses has little impact on the structure of the maps and clusters; in fact, they stay the same. However, changes are significant for the amount of citations gathered by some authors, which also affects which author names and documents get included in the clusters.

The CAD analysis is also very useful for delimiting the document set and for concentrating on a certain period and research area. However, the restrictions enforced in the CAD analyses have drawbacks as well. Not only does the CAD analysis eliminate material published in other fields of research, but also LIS research published in media not covered in the ISI databases. This excludes research published in non-English journals, books and not the least, Web documents. This may lead to a risk of missing emerging research areas and issues. A study on XML (extensible markup language) research suggests a low interaction between research areas publishing in journals and on the Web, emphasizing the risk of not identifying important research issues by focusing the analysis only on either ISI covered journals or on Web documents (Zhao, 2003). In the XML research case, examples on research primarily published in Web documents concerns issues on programming and processing of data, as well as research on the "semantic Web," research issues of relevance in LIS, which might be more visible if Web documents were also analyzed.

Acknowledgments

The author wishes to express his gratitude to Professor Olle Persson for supervision throughout this work; the anonymous reviewers for their valuable comments and helpful suggestions; Professors Bluma Peritz, Katherine `McCain, and Mike Thelwall for valuable comments on an early draft of this article at the ISSI 2005 Doctoral Forum; and the Nordic Research School in LIS (NoRSLIS) for providing an invaluable research environment and partly funding the author's attendance at ISSI 2005 (The 10th International Conference of the International Society for Scientometrics and Informetrics).

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: CoLIS4 (pp. 185–197). Westport, CT: Libraries Unlimited.

DOI: 10.1002/asi

Björneborn, L., & Ingwersen, P. (2001). Perspective of webometrics. Scientometrics, 50(1), 65–82.

Burrell, Q.L. (2003). Predicting future citation behavior. Journal of the American Society for Information Science and Technology, 54, 372–378.

Chen, C.M. (2004). Searching for intellectual turning points: Progressive knowledge domain visualization. Proceedings of the National Academy of Sciences of the United States of America, 101, 5303–5310.

Feitelson, D.G., & Yovel, U. (2003). Predictive ranking of computer scientists using CiteSeer data. Journal of Documentation, 60(1), 44–61.

Garfield, E. (1994). Research fronts. Current Contents, 41, 3–7.

Glänzel, W. (1997). On the possibility and reliability of predictions based on stochastic citation processes. Scientometrics, 40, 481–492.

Glänzel, W., & Schoepflin, U. (1994). A stochastic model for the aging of scientific literature. Scientometrics, 30(1), 49–64.

Glänzel, W., & Schubert, A. (1995). Predictive aspects of a stochastic model for citation processes. Information Processing & Management, 31(1), 69–80.

Glänzel, W., Thijs, B., & Schlemmer, B. (2004). A bibliometric approach to the role of author self-citations in scientific communication. Scientometrics, 59(1), 63–77.

Ingwersen, P., & Järvelin, K. (2005). The turn: integration of information seeking and retrieval in context. Dordrecht: Springer.

Morris, S.A. (2004). Research topics to reference author crossmap: Information science, 12 journals: Paper group by reference author crossmap. Retrieved January 5, 2006, from http://samorris.ceat.okstate.edu/web/info_sci/default.htm

Morris, S.A., Yen, G., Wu, Z., & Asnake, B. (2003). Time line visualization of research fronts. Journal of the American Society for Information Science and Technology, 54, 413–422.

Morris, T.A., & McCain, K.W. (1998). The structure of medical informatics journal literature. Journal of the American Medical Informatics Association, 5, 448–466.

Persson, O. (1994). The intellectual base and research fronts of 1986–1990. Journal of the American Society for Information Science, 45(1), 31–38.

Persson, O. (2001). All author co-citations versus first author co-citations. Scientometrics, 50, 339–344.

Persson, O. (2005). Exploring the analytical potential of comparing citing and cited source items. In P. Ingwersen & B. Larsen (Eds.), Proceedings of ISSI 2005: The 10th International Conference of the International Society for Scientometrics and Informetrics (pp. 24–33). Stockholm: Karolinska University Press.

Persson, O. (2006). Bibexcel: A toolbox for bibliometricians developed by Olle Persson (Version 2006–01–03). Retrieved January 5, 2006, from http://www.umu.se/inforsk/Bibexcel/index.html

Price, D.J.D. (1965). Networks of scientific papers. Science, 149(3683), 510–515.

Saracevic, T. (1999). Information science. Journal of the American Society for Information Science, 50, 1051–1063.

Thomson ISI. (2003). Journal citation reports: JCR social sciences. Philadelphia: Author.

Thomson ISI. (2005). Web of Science. Retrieved August 7–10, 2005, from http://portal.isiknowledge.com/portal.cgi/wos

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, 163–172.

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

Yashin, Y.I., & Yashin, A.Y. (2001). Current trends in gas chromatography methods and instrumentation: A scientometric study. Journal of Analytical Chemistry, 56, 200–213.

Zhao, D. (2003). A comparative citation analysis study of web-based and print-journal based scholarly communication in the XML research field. Unpublished doctoral dissertation, The Florida State University, Tallahassee. Retrieved June 14, 2006, from http://etd.lib.fsu.edu/theses/available/etd-09232003-012028/

DOI: 10.1002/asi