Impact of Data Sources on Citation Counts and Rankings of LIS Faculty: Web of Science Versus Scopus and Google Scholar

Lokman I. Meho and Kiduk Yang
School of Library and Information Science, Indiana University, Bloomington, IN 47405.
E-mail: meho@indiana.edu; kiyang@indiana.edu

The Institute for Scientific Information's (ISI, now Thomson Scientific, Philadelphia, PA) citation databases have been used for decades as a starting point and often as the only tools for locating citations and/or conducting citation analyses. The ISI databases (or Web of Science [WoS]), however, may no longer be sufficient because new databases and tools that allow citation searching are now available. Using citations to the work of 25 library and information science (LIS) faculty members as a case study, the authors examine the effects of using Scopus and Google Scholar (GS) on the citation counts and rankings of scholars as measured by WoS. Overall, more than 10,000 citing and purportedly citing documents were examined. Results show that Scopus significantly alters the relative ranking of those scholars that appear in the middle of the rankings and that GS stands out in its coverage of conference proceedings as well as international, non-English language journals. The use of Scopus and GS, in addition to WoS, helps reveal a more accurate and comprehensive picture of the scholarly impact of authors. The WoS data took about 100 hours of collecting and processing time, Scopus consumed 200 hours, and GS a grueling 3,000 hours.

Introduction

Academic institutions, federal agencies, publishers, editors, authors, and librarians increasingly rely on citation analysis, along with publications assessment and expert opinions, for making hiring, promotion, tenure, funding, and/or reviewer and journal evaluation and selection decisions. In general, citation counts or rankings are considered partial indicators of research impact and quality, often used to support or question other indicators such as peer judgment (Borgman & Furner, 2002; Cronin, 1984; Holden, Rosenberg, & Barker, 2005; Moed, 2005; van Raan, 1996, 2005; Wallin, 2005).

Many scholars have argued for and against the use of citations for assessing research impact or quality. Proponents have reported the validity and reliability of citation counts in research assessments as well as the positive correlation between these counts and peer reviews and assessments of publication venues (Aksnes & Taxt, 2004; Glänzel, 1996; Kostoff, 1996; Martin, 1996; Narin, 1976; So, 1998; van Raan, 2000). Critics, on the other hand, claim that citation counting has serious problems or limitations that affect its validity (MacRoberts & MacRoberts, 1996; Seglen, 1998). Important limitations reported in the literature focus on, among other things, the problems associated with the data sources used, especially the Institute for Scientific Information (ISI; currently Thomson Scientific, Philadelphia, PA) citation databases: Arts & Humanities Citation Index, Science Citation Index, and Social Sciences Citation Index—the standard and most widely used tools for generating citation data for research and other assessment purposes. These tools are now currently part of what is known as *Web of Science* (WoS), the portal used to search the three ISI citation databases. In this article, we use ISI citation databases and WoS interchangeably.

Critics of ISI citation databases note that they (a) cover mainly North American, Western European, and English-language titles; (b) are limited to citations from 8,700 journals;[1] (c) do not count citations from books and most conference proceedings; (d) provide different coverage between research fields; and (e) have citing errors, such as homonyms, synonyms, and inconsistency in the use of initials and in the spelling of non-English names (many of these errors, however, come from the primary documents themselves rather than being the result of faulty ISI indexing).

[1]Ulrich's Periodicals Directory (Bowker, New Providence, NJ) lists approximately 22,500 active academic/scholarly, refereed journals. Of these, approximately 7,500 are published in the United States, 4,150 in the United Kingdom, 1,600 in the Netherlands, 1,370 in Germany, 660 in Australia, 540 in Japan, and 500 in Canada, 450 in China, 440 in India, and 420 in France.

Received October 16, 2006; revised December 19, 2006; accepted January 22, 2007

 • Published online 30 August 2007 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/asi.20677

Studies that have addressed problems of, and/or suggested alternative or complementary sources to, ISI citation databases are very few and can be divided into two main groups:

1. Studies that examined the effect of certain limitations in the ISI database, most often comparing its coverage with that of other citation sources
2. Studies that suggested or explored different or additional sources and methods for identifying citations

Studies That Examined the Effect of Coverage Limitations in ISI Databases

In a study aimed at analyzing the effect of the omission of certain journals in ISI databases on citation-based appraisals of communication literature, Funkhouser (1996) examined references in 27 communication journals (13 covered by ISI and 14 not covered) for the year of 1990. He found that 26% of author citations were from non-ISI journals and that 27 of the 50 most highly cited authors received at least 25% of their citations from non-ISI journals. Funkhouser, however, did not verify whether the omission of those 14 journals had any impact on the relative citation ranking of scholars if one relied only on ISI data.

Cronin, Snyder, and Atkins (1997) analyzed thousands of references from monographs and leading academic journals in sociology to identify the effects of ISI databases' noncoverage of citations in monographic literature. They found that the relative rankings of authors who were highly cited in the monographic literature did not change in the journal literature of the same period. The overlap of citations in monographs and journals, however, was small, suggesting that there may be two distinct populations of highly cited authors.

Whitley (2002) compared the duplication and uniqueness of citing documents in Chemical Abstracts and Science Citation Index for the works of 30 chemistry researchers for the years 1999–2001. She found that 23% of all the citing documents were unique to Chemical Abstracts, 17% were unique to the Science Citation Index, and the remaining 60% were duplicated in the two databases. Whitley concluded that relying on either index alone would lead to faulty results when trying to obtain citation totals for individual authors.

Goodrum, McCain, Lawrence, and Giles (2001) and Zhao and Logan (2002) compared citations from CiteSeer/ResearchIndex, a Web-based citation indexing system, with those from ISI's Science Citation Index (SCI) in the field of computer science. Both studies found a 44.0% overlap among the top-25 cited authors and concluded that CiteSeer/ResearchIndex and SCI were complementary in their coverage of the field. Recently, Pauly and Stergiou (2005) compared citation counts between WoS and GS for papers in mathematics, chemistry, physics, computing sciences, molecular biology, ecology, fisheries, oceanography, geosciences, economics, and psychology. Each discipline was represented by three authors, and each author was represented by three articles (i.e., 99 articles in total). The authors also examined citations to an additional 15 articles for a total of 114. Without assessing the accuracy or relevance and quality of the citing articles, the authors reported such good correlation in citation counts between the two sources that they suggested GS can substitute for WoS.

Bauer and Bakkalbasi (2005) compared citation counts provided by WoS, Scopus, and Google Scholar (GS) for articles from the *Journal of the American Society for Information Science and Technology* published in 1985 and in 2000. They found that WoS provided the highest citation counts for the 1985 articles and GS provided statistically significant higher citation counts than either WoS or Scopus for the 2000 articles. They did not find significant differences between WoS and Scopus for either year. The authors, however, stated that more rigorous studies were required before these findings could be considered definitive, especially because the scholarly value of some of the unique material found in GS remained an open question.

Jacsó (2005a) also conducted several tests comparing GS, Scopus, and WoS, searching for documents citing (a) Eugene Garfield, (b) an article by Garfield published in 1955 in *Science*, (c) the journal *Current Science*, and (d) the 30 most-cited articles from *Current Science*. He found that coverage of *Current Science* by GS is "abysmal" and that there is considerable overlap between WoS and Scopus. He also found many unique documents in each source, claiming that the majority of the unique items were relevant and substantial. For lack of space, Jacsó's analysis was limited to reporting citation counts and retrieval performance by time period; he did not provide an in-depth analysis and examination of, for example, the type, refereed status, and source of the citations.

Bakkalbasi, Bauer, Glover, and Wang (2006) compared citation counts for articles from two disciplines (oncology and condensed matter physics) and 2 years (1993 and 2003) to test the hypothesis that the different scholarly publication coverage provided by WoS, Scopus, and GS would lead to different citation counts from each. They found that for oncology in 1993, WoS returned the highest average number of citations; 45.3, Scopus returned the highest average number of citations (8.9) for oncology in 2003; and WoS returned the highest number of citations for condensed matter physics in 1993 and 2003 (22.5 and 3.9, respectively). Their data showed a significant difference in the mean citation rates between all pairs of resources except between Scopus and GS for condensed matter physics in 2003. For articles published in 2003, WoS returned the largest amount of unique citing material for condensed matter physics and GS returned the most for oncology. Bakkalbasi, Bauer, Glover, and Wang concluded that all three tools returned some unique material and that the question of which tool provided the most complete set of citing literature might depend on the subject and publication year of a given article.

Studies That Suggested Sources and Methods Beyond ISI or Citation Databases

In a 1995 article, Reed recommended that faculty seeking tenure or promotion: (a) review the citations in selected key

DOI: 10.1002/asi

journals in their specialty that were not covered in ISI databases; (b) scan the citations and bibliographies in textbooks and monographs pertinent to their research areas; (c) record citations discovered through research, teaching activities, and professional reading throughout their careers; and (d) maintain a continuously updated file of citations as they are discovered. These recommendations were adopted by Nisonger (2004a), who additionally suggested that sources be examined (e.g., books, journal articles, and doctoral dissertations identified in major bibliographies in one's specialty area), and that the author's name be searched on the Web for items not indexed in ISI databases.

Unlike Reed who only compiled and recommended a list of techniques to locate citations not covered by ISI, Nisonger (2004a) conducted a self-study to show how ISI coverage compared to citation data he collected using the aforementioned six techniques. His study was based on analysis of his own lifetime citation record, which he compiled by (a) searching the ISI databases, (b) manually searching the literature for nearly 15 years, and (c) making use of various Web search engines. He found that (with self-citations excluded) ISI captured 28.8% of his total citations, 42.2% of print citations, 20.3% of citations from outside the United States, and 2.3% of non-English citations. Nisonger suggested that faculty should not rely solely on ISI author citation counts, especially when demonstration of international impact is important. He also suggested that rankings based on ISI data of a discipline's most-cited authors or academic departments might be significantly different if non-ISI citation data were included. This suggestion, however, was not verified by empirical data; it merely suggested that broader sourcing of citations might alter one's relative ranking vis-à-vis others.

Emergence of Competitors to Web of Science

Both Reed's recommendations and Nisonger's methods are useful techniques for locating citations; however, they are not practical in the case of large study samples. Citation databases remain the most viable methods for generating bibliometric data and for making accurate citation-based research assessments and large-scale comparisons between works, authors, journals, and departments. Until recently, WoS was the standard tool for conducting extensive citation searching and bibliometric analysis, primarily because it was the only general and comprehensive citation database in existence. This, however, may no longer be the case because several databases or tools that provide citation searching capabilities have appeared in the past few years. These databases or tools, which currently number over 100, can be classified into three basic categories. The first allows the user to search in the full text field to determine whether certain papers, books, authors, or journals have been cited in a document. Examples of these databases or tools include ACM (Association for Computing Machinery) Digital Library, arXiv.org, Emerald Full Text, ERIC, Google Book Search, IEEE (Institute of Electrical and Electronics Engineers) Computer Society Digital Library and IEEE Xplore, Library Literature and Information Science Full Text, NetLibrary, and Elsevier's Scirus. Also belonging to this category are databases or tools that automatically extract and parse bibliographic information and cited references from electronic fulltext documents retrieved from personal homepages and digital archives and repositories. Examples of these include CiteSeer (computer science), Google Scholar (general), RePEc (economics), and SMEALSearch (business).

The second category of databases or tools allows the user to search in the cited references field to identify relevant citations. Examples of these include several of EBSCO's products (e.g., Academic Search Premier and Library, Information Science & Technology Abstracts), PsycINFO, PubMed Central, and Elsevier's ScienceDirect. The last category includes databases that serve exactly like WoS (i.e., those designed primarily for citation searching, but used for bibliographic searching too); the main and perhaps only good example of this category is Scopus.

Details about the citation searching features and strengths and weaknesses of the aforementioned and many other databases that allow citation searching can be found in Roth (2005), Ballard and Henry (2006), and the many review and scholarly articles by Péter Jacsó (http://www2.hawaii.edu/~jacso/).

Research Questions and Significance

The emergence of new citation databases or sources, especially those that are comprehensive and/or multidisciplinary in nature (e.g., Scopus and Google Scholar), pose a direct challenge to the dominance of WoS and raise questions about the accuracy of using it exclusively in citation, bibliometric, and scholarly communication studies. Thus, several questions suggest themselves for future studies:

1. What is the impact of using new, additional citation databases or tools on the counting and ranking of works, authors, journals, and academic departments?
2. How do the citations generated by these new sources compare with those found in WoS in terms of, for example, document source, document type, refereed status, language, and subject coverage?
3. Do these new citation sources represent alternatives or complements to WoS?
4. What strengths and weaknesses do these new citation sources have relative to WoS?

Answering these questions is important to anyone trying to determine whether an article, author, or journal citation search should be limited to WoS. The answers to these questions are also important for those seeking to use appropriate tools to generate more precise citation counts, rankings, and assessments of research impact than those based exclusively on WoS. More complete citation counts can help support or identify more precisely discrepancies between research productivity, peer evaluation, and citation data. More complete citation counts can also help generate more accurate h-index

DOI: 10.1002/asi

scores for authors and journals (Bar-Ilan, 2006; Bornmann & Daniel, 2005, 2007; Cronin & Meho, 2006; Hirsch, 2005; Oppenheim, 2007) and journal impact factors (Garfield, 1996, 2006; Nisonger, 2004b; Saha, Saint, & Christakis, 2003), as well as identify international impact (Nisonger, 2004a; de Arenas, Castanos-Lomnitz, & Arenas-Licea, 2002).

Scholars trying to locate citations to a specific publication for traditional research purposes (as opposed to citation counts, research evaluation, and so on) will find answers to the aforementioned questions very useful, too, especially in cases where bibliographic searches fail to identify relevant research materials. Serials librarians who use citation counts and analyses to make journal subscription and cancellation decisions will benefit from studies addressing these questions as well because the findings of such studies will show whether there is a need to rely on multiple sources of citation data. Vendors and producers of fulltext databases, such as ProQuest-CSA (Ann Arbor, MI/Bethesda, MD), EBSCO (EBSCO Information Services, Birmingham, AL), Elsevier (The Netherlands), OCLC (Online Computer Library Center, Dublin, OH), Ovid (Ovid Technologies, New York, NY), Sage (Thousand Oaks, CA), Springer (Berlin/Heidelberg/New York), Taylor & Francis (London/ Philadelphia), and H. W. Wilson (Bronx, NY) will also benefit from answers to these questions by applying the findings to develop and illustrate additional features and uses of their databases.

Although several authors have attempted to answer the aforementioned questions (see studies reviewed above), these authors agree that more research is required before reaching definitive conclusions about, among other things, the effects of using multiple citation sources on the citation counts and rankings of scholars. The current study builds on these previous attempts by:

- Analyzing the effects of using Scopus and GS on the citation counts and rankings of individual scholars as measured by WoS, using citations to the work of 25 library and information science (LIS) faculty members as a case study. These faculty members make an ideal case study due to the interdisciplinary and multidisciplinary nature of their research areas and their use of, and reliance on, various types of literature for scholarly communication (e.g., journal articles, conference papers, and books).
- Examining the similarities and differences between WoS, Scopus, and GS in terms of coverage period, sources of citations, document type, refereed status, language, and subject coverage, and identifying strengths and weaknesses of the three tools.
- Discussing the implications of the findings on citation analysis and bibliometric studies.

Scopus and GS were chosen because of their similarity to WoS in that they were created specifically for citation searching and bibliometric analysis, in addition to being useful for bibliographic searches. Scopus and GS were also chosen because they represent the only real or potential competitors to WoS in citation analysis and bibliometrics research areas. More information about these three sources is provided below.

METHODS

Citation Databases or Tools

WoS, which comprises the three ISI citation databases (Arts & Humanities Citation Index, Science Citation Index, and Social Sciences Citation Index), has been the standard tool for a significant portion of citation studies worldwide. A simple keyword search in WoS and other databases (e.g., Library and Information Science Abstracts, Pascal, Medline, EMBASE, Biosis Previews, and INSPEC) indicates that ISI databases have been used, or referred to, in several thousand journal articles, conference papers, and chapters in books published in the last three decades. WoS's Web site provides substantial factual information about the database, including the number of records and the list of titles indexed. It also offers powerful features for browsing, searching, sorting, and saving functions, as well as exporting to citation management software (e.g., EndNote [Thomson ResearchSoft, Philadelphia, PA] and RefWorks [Bethesda, MD]). Coverage in WoS goes back to 1900 for Science Citation Index, 1956 for Social Sciences Citation Index, and 1975 for Arts & Humanities Citation Index. As of October 2006, there were over 36 million records in the database (the version the authors had access to) from approximately 8,700 scholarly titles (Thomson Corporation, 2006a), including several hundred conference proceedings and over 190 open access journals (Harnad & Brody, 2004).[2] Over 100 subjects are covered in WoS, including all the major arts, humanities, sciences, and social sciences subdisciplines (e.g., architecture, biology, business, chemistry, health sciences, history, medicine, political science, philosophy, physics, religion, and sociology). For more details on WoS, see Goodman and Deis (2005) and Jacsó (2005a). Similar to ISI, Elsevier, the producer of Scopus, provides substantial factual information about the database, including the number of records and the list of titles indexed. It also offers powerful features for browsing, searching, sorting, and saving functions, as well as exporting to citation management software. Coverage in Scopus goes back to 1966 for bibliographic records and abstracts and 1996 for citations. As of October 2006, there were over 28 million records in the database from over 15,000 peer-reviewed titles, including coverage of 500 Open Access journals, 700 conference proceedings, 600 trade publications, and 125 book series (Elsevier science Publishers, 2006). Subject areas covered in Scopus include Chemistry, Physics, Mathematics, and Engineering (4,500 titles), Life and Health Sciences (5,900 titles, including 100% Medline

[2]The figure for conference proceedings was generated by analyzing the source titles of over 125,000 records that were published in the Lecture Notes series (e.g., *Lecture Notes in Artificial Intelligence*, *Lecture Notes in Computer Science*, and *Lecture Notes in Mathematics*). Also analyzed were the indexed titles that included the word conference, proceedings, symposium, workshop, or meeting in their names.

DOI: 10.1002/asi

coverage), Arts and Humanities, Social Sciences, Psychology, and Economics (2,700 titles), Biological, Agricultural, and Environmental Sciences (2,500 titles), and General Sciences (50 titles). For more details on Scopus, see Goodman and Deis (2005) and Jacsó (2005a).

In contrast to ISI and Elsevier, Google does not offer a publisher list, title list, document type identification, or any information about the time span or the refereed status of records in GS. This study, however, found that GS covers print and electronic journals, conference proceedings, books, theses, dissertations, preprints, abstracts, and technical reports available from major academic publishers, distributors, aggregators, professional societies, government agencies, and preprint/reprint repositories at universities, as well as those available across the web. Examples of these sources include: Annual Reviews, arXiv.org, ACM, Blackwell, Cambridge Scientific Abstracts (CSA), Emerald, HighWire Press, Ingenta, IEEE, PubMed, Sage, Springer, Taylor & Francis, University of Chicago Press, and Wiley, among others (Bauer & Bakkalbasi, 2005; Gardner & Eng, 2005; Jacsó, 2005b; Noruzi, 2005; Wleklinski, 2005). Although GS does not cover material from all major publishers (e.g., American Chemical Society and Elsevier), it identifies citations to articles from these publishers when documents from other sources cite these articles. Google Scholar does not indicate how many documents it searches.

Table 1 provides detailed information about the breadth and depth of coverage, subject coverage, citation browsing and searching options, analytical tools, and downloading and exporting options of all three sources.

Units of Analysis

To analyze the effect of using additional sources to WoS on the citation counts and rankings of LIS faculty members and to be able to generalize the findings to the field, we explored the difference Scopus and GS make to results from WoS for all 15 faculty members of the School of Library and

TABLE 1. Comparisons of databases and tools used in the study.

	Web of Science	Scopus	Google Scholar
Breadth of coverage	36 million records (1955-) 8,700 titles (including 190 open access journals and several hundred conference proceedings)	28 million records (1966-) 15,000 titles (including 12,850 journals, 700 conference proceedings, 600 trade publications, 500 open access journals, and 125 book series)	Unknown number of records Unknown number of sources Over 30 different document types Unknown number of publishers
Depth of coverage	A&HCI: 1975- SCI: 1900- SSCI: 1956-	With cited references data: 1996- Without cited references data: 1966-	Unknown
Subject coverage	All	All	All
Citation browsing options	Cited author Cited work	Not available	Not available
Citation searching options	Cited author Cited work (requires use of the abbreviated journal, book, or conference title in which the work appeared) Cited year	The “Basic Search” interface allows keyword and phrase searching via “References” field. The “Advanced Search” interface, allows searching for: Cited author (REFAUTH) Cited title (REFTITLE) Cited work (REFSRCTITLE) Cited year (REFPUBYEAR) Cited page (REFPAGE) Cited reference (REF), which is a combined field that searches the REFAUTH, REFTITLE, REFSRCTITLE, REFPUBYEAR, and REFPAGE fields	Keyword and phrase searching Limit/search options include “Author,” “Publication,” “Date,” and “Subject Areas”
Analytical tools	Ranking by author, publication year, source name, country, institution name, subject category, language, and document type	Ranking by author, publication year, source name, subject category, and document type Analysis of citations by year (via Citation Tracker)	Not available
Downloading and exporting options to citation management software (e.g., EndNote and RefWorks)	Yes	Yes	Yes

Note. A&HCI = Arts & Humanities Citation Index; SCI = Science Citation Index; SSCI = Social Science Citation Index.

DOI: 10.1002/asi

Information Science at Indiana University-Bloomington (SLIS).[3] These faculty members cover most of the mainstream LIS research areas as identified by the Association of Library and Information Science Education (ALISE, 2006); they also cover research areas beyond those listed by ALISE (e.g., computer-mediated communication and computational linguistics). Moreover, SLIS faculty members are the most published and belong to one of the most cited American Library Association-accredited LIS programs in North America (Adkins & Budd, 2006; Budd, 2000; Persson & Åström, 2005; Thomson Corporation, 2006b). From 1970 to December 2005, the 15 SLIS faculty members had published or produced over 1,093 scholarly works, including 312 refereed journal articles, 305 conference papers (almost all refereed), 131 chapters (some refereed), 93 nonrefereed journal articles, 83 technical reports or working papers, 59 articles in professional journals, 36 books, 35 edited volumes, and 12 refereed review articles, among others. The citations to the work of an additional 10 faculty members were examined in the study to verify the findings and conclusions made based on data for the main study group (see more below).

All data were entered into EndNote libraries and Access databases and were coded by citing source (e.g., journal name, conference, book, and so on), document type (e.g., journal article, review article, conference paper, and so on), refereed status of the citing item, year, language, and source used to identify the citation. The refereed status of the sources of citations was determined through Ulrich's International Periodicals Directory and the domain knowledge of the researchers and their colleagues.

Data Collection

All WoS and Scopus data were manually collected and processed twice by one of the authors (LIM) in October 2005 and again (for accuracy and updating purposes) in March 2006. The GS data were harvested in March 2006; however, identifying their relevancy and full bibliographic information took approximately 3,000 hours of work over a 6-month period, which included manually verifying, cleaning, formatting, standardizing, and entering the data into EndNote Libraries and Access databases.[4]

The Cited Author search option was used in WoS to identify citations to each of the 1,093 items published by the 15 faculty members constituting the main study group and to the 364 items published by the 10 faculty members constituting the test group. Citations to items in which the faculty members were not first authors, as well as citations to dissertations and other research materials written by them, were included in the study. Although publicly available, the data have been made anonymous, assigning citations to faculty members by their research areas rather than by names.

Unlike WoS, Scopus does not have browsing capabilities for the cited authors or cited works fields that would allow limiting the search to relevant citations (cited works field is the index field for names of journals, books, patent numbers, and other publications). As a result, instead of browsing the cited authors or cited works fields, we used an exact match search approach to identify all potentially relevant citations in the database. This method uses the title of an item as a search statement (e.g., Invoked on the Web) and tries to locate an exact match in the cited references field of a record. Using the titles of the 1,457 items published or produced by the 25 faculty members included in the study, the method allowed us to identify the majority of the relevant citations in the database. In cases where the title was too short or ambiguous to refer to the item in question, we used additional information as keywords (the first author's last name and, if necessary, journal name and/or book or conference title) to ensure that we retrieved only relevant citations. In cases where the title was too long, we used the first few words of the title because utilizing all the words in a long title increases the possibility of missing some relevant citations due to typing or indexing errors. When in doubt, we manually examined all retrieved records to make sure they cited the items in question. Other search options in the database were used (e.g., Author Search and Advanced Search), but they not only did not identify any unique, additional citations, they were less inclusive than the exact match approach. For example, because not all of the 1,457 items published by the 25 faculty members are indexed in the database, the Author Search approach would have been inappropriate or would have resulted in incomplete sets of relevant citations.

Google Scholar was searched for citations using two methods: author search and exact match (or exact phrase) search. The author search usually retrieves items published by an author and ranks the items in a rather inconsistent way. Once the items are retrieved, a click on the "Cited by . . ." link allows the searcher to display the list of citing documents. The Cited-by link is automatically generated by GS for each cited item.

The exact match search approach was used to ensure that citations were not missed due to errors in GS's author search algorithm. This search strategy, which is the same as applied in Scopus, resulted in 1,301 records. Of these, 534 were unique relevant citations. In other words, if the exact match search approach was not used along with the author search approach, 534 (or 14.6%) of GS's relevant citations would have been missed. The remaining 767 records retrieved through the exact match search were either previously found through the author search approach or were not relevant. Almost all of the false drops were documents retrieved when searching for citations to short-title items.

A major disadvantage of GS is that its records are retrieved in a way that is very impractical for use with large sample

[3]As of January 2007, SLIS had 17 full-time faculty members.

[4]At the time of data collection, GS did not provide the option of downloading search results into a bibliographic management software program (e.g., EndNote, BibTeX (Open Directory Project), and RefWorks). Although the ability to download search results into any of these programs would have reduced the amount of time spent on processing the citations, manually verifying, cleaning, formatting, and standardizing the citations would still have been necessary and would have consumed an excessive amount of time.

DOI: 10.1002/asi

sizes, requiring a very tedious process of manually extracting, verifying, cleaning, organizing, classifying, and saving the bibliographic information into meaningful and usable formats. Moreover, unlike WoS and Scopus, GS does not allow resorting of the retrieved sets in any way (such as by date, author name, or data source); as mentioned earlier, retrieved records in GS are rank-ordered in a rather inconsistent way. The result sets show short entries, displaying the title of the cited article and the name of the author(s) and, in some cases, the source. Entries that include the link Cited by indicate the number of times the article has been cited. Clicking on this link will take users to a list of citing articles. Users will be able to view the fulltext of only those items that are available for free and those to which their libraries subscribe.

Other major disadvantages of GS include duplicate citations (e.g., counting a citation published in two different forms, such as preprint and journal article, as two citations), inflated citation counts due to the inclusion of nonscholarly sources (e.g., course reading lists), phantom or false citations due to the frequent inability of GS to recognize real matches between cited and citing items claiming a match where there is not even minimal "chemistry" (Jacsó, 2006), errors in bibliographic information (e.g., wrong year of publication), as well as the lack of information about document type, document language, document length, and the refereed status of the retrieved citations. In many cases, especially when applying the Exact Phrase search method, the item for which citations are sought is retrieved and considered a citation by GS (in such cases, these citations were excluded from the search results). Perhaps the most important factor that makes GS very cumbersome to use, is the lack of full bibliographic information for citations found. Even when some bibliographic information is made available (e.g., source), it is not provided in a standard way thus requiring a considerable amount of manual authority control, especially among citations in conference proceedings. For example, the annual meeting of the American Society for Information Science and Technology is cited in at least five different ways (ASIST 2004: ..., ASIST 2005: ..., Proceedings of the American Society for Information Science and Technology, Annual Meeting of the ..., and so on), whereas in WoS and Scopus almost all entries for this conference and other conference proceedings are entered in a standardized fashion. The presence of all these problems in GS suggests that unless a system is developed that automatically and accurately parses result sets into error-free, meaningful, and usable data, GS will be of limited use for large-scale comparative citation and bibliometric analyses.

To make sure that citations were not overlooked because of searching or indexing errors, we looked for the bibliographic records of all citations that were missed by one or two of the three tools. For example, if a citation was found in WoS, but not in Scopus or GS, we conducted bibliographic searches in Scopus and/or GS to see if the item were in fact indexed in them. When the bibliographic record of any of these missed citations was found in one of the three tools, we examined (a) why it was not retrieved through the citation search methods described above, and (b) whether it should be counted as a citation. Items that were overlooked due to searching errors (16 in the case of WoS and 27 in Scopus) were counted as citations toward their respective databases; most of the searching errors were due to having missed selecting a relevant entry when browsing the cited references field in WoS and making typographical errors when entering a search query in Scopus. Items that were missed due to database/system errors were tallied, but were not counted as citations. These included:

- WoS: Ten citations were missed due to incomplete lists of references. These citations are the equivalent of 0.5% of the database's relevant citations.
- Scopus: Seventy-five citations were missed due to lack of cited references information and 26 citations due to incomplete lists of references in their respective records. In total, Scopus missed 101 (or the equivalent of 4.4% of its relevant citations) due to database errors.
- GS: Missed 501 (or the equivalent of 12.0% of its relevant citations) due to system errors. Many of the errors in GS were a result of matching errors. For example, the search engine failed, in many cases, to identify an exact match with the search statements used because a word or more in the title of the cited item was automatically hyphenated in the citing document. Or GS failed to retrieve relevant citations from documents that do not include well-defined sections such as Bibliography, Cited References, Cited Works, Endnotes, Footnotes, or References.

These results suggest that if citation searching of individual LIS scholars were limited to Scopus, a searcher would miss an average of 4.4% of the relevant citations due to database errors. In the case of GS, the percentage would be 12.0%; this percentage would increase to 26.6% had we not used the Exact Phrase search approach described earlier. The results also suggest that when using GS one must use both the Author search and Exact Phrase search methods.

It is important to note here that it took about 100 hours of work to collect, clean, standardize, and enter all the data into EndNote libraries and Access databases from WoS, about 200 hours in the case of Scopus, and, as mentioned earlier, over 3,000 hours in the case of GS. In other words, collecting GS data took 30 as much time as collecting WoS data and 15 as much time as that of Scopus—this includes the time needed to double-check the missed items in each source. It is also important to note that in studies such as this, it is essential that the investigators have access to complete lists of publications of the authors being examined. Without this information, there would be major problems with the data collected, especially when there are authors with common names among the study sample. In our case, all 25 faculty members constituting the study and test groups had their complete publication information available online or they provided it on request. This information was very useful in the case of approximately half of the faculty members as we discovered multiple authors with the names B. Cronin,

DOI: 10.1002/asi

S. Herring, J. Mostafa, N. Hara, D. Shaw, and K. Yang. The availability of their publication lists helped avoid including nonrelevant citations.

RESULTS AND DISCUSSION

The results of this study are presented and discussed in three sections: (a) the effect of using Scopus on the citation counts and rankings of the 15 SLIS faculty members as measured by WoS; (b) the effect of using GS on the citation counts and rankings as measured by WoS and Scopus combined; and (c) the sources of citations found in all three tools, including their names (i.e., journal and conference proceedings), refereed status, and language. The results of the test group are discussed where needed (see below). Because the three tools provide different citation coverage in terms of document type and time period, we limited most of the analysis to citations from types of documents and years common to all three tools, that is, conference papers and journal items (e.g., journal articles, review articles, editorials, book reviews, and letters to the editor) published between 1996 and 2005. Excluded from the analysis are citations found in books, dissertations, theses, reports, and so on, as well as 475 citations from GS that did not have complete bibliographic information. These 475 citations primarily included bachelor's theses, presentations, grant and research proposals, doctoral qualifying examinations, submitted manuscripts, syllabi, term papers, working papers, Web documents, preprints, and student portfolios.

Effect of Scopus on Citation Counts and Rankings of SLIS Faculty

To show the difference that Scopus makes to the citation counts and rankings of SLIS faculty members as measured by WoS, we compare the number of citations retrieved by both databases, show the increase Scopus makes toward the total number of citations of SLIS as a whole and also of individual faculty members, and explore the effect Scopus has on altering the relative citation ranking of SLIS faculty members. We also examine the overlap and unique coverage between the two databases. The refereed status of citations found in WoS and Scopus is not discussed because the great majority of citations from these two databases come from scholarly, peer-reviewed journals and conference proceedings.

As shown in Tables 2 and 3, Scopus includes 278 (or 13.7%) more citations than WoS, suggesting that Scopus provides more comprehensive coverage of the SLIS literature than WoS.[5] Further analysis of the data shows that combining citations from Scopus and WoS increases the number of citations of SLIS as a whole by 35.1% (from 2,023 to 2,733 citations). This means that if only WoS was used to locate citations for SLIS faculty members, on average, more than one third of relevant citations (found in the union of WoS and Scopus) would be missed; the percentage of missed citations would be 18.8% were only Scopus used.

TABLE 2. Citation count by year-Web of Science and Scopus

Years	WoS	Scopus	Union of WoS and Scopus
1971–1975	1	–	1
1976–1980	15	–	15
1981–1985	129	–	129
1986–1990	201	–	201
1991–1995	323	–	323
Subtotal	669	–	669
1996	119	101	140
1997	121	119	144
1998	142	123	167
1999	131	128	164
2000	175	171	219
2001	207	242	278
2002	202	220	271
2003	251	291	348
2004	323	459	510
2005	352	447	492
Subtotal	2,023	2,301	2,733
Total	2,692	2,301	3,402

Perhaps more importantly, the data show that the percentage of increase in citation counts for individual faculty members varies considerably depending on their research areas, ranging from 4.9% to 98.9%. For example, faculty members with research strengths in such areas as communities of practice, computational linguistics, computer-mediated communication, data mining, data modeling, discourse analysis, gender and information technology, human–computer interaction, information retrieval, information visualization, intelligent interfaces, knowledge discovery, and user modeling, will find their citation counts increase considerably more than those faculty members with research strengths in other areas (see Table 3). These findings not only imply that certain subject areas will benefit more than others from using both Scopus and WoS to identify relevant citations, they also suggest that to generate accurate citation counts for faculty members, and by extension schools, and to accurately compare them to one another, a researcher must use both databases. The importance of using Scopus in addition to WoS is further evidenced by:

- The relative ranking of faculty members changes in 8 out of 15 cases, strikingly so in the cases of faculty members E, F, H, and I (see Table 4). Although the overall relative ranking of the faculty members does not change significantly when citations from both databases are counted (Spearman rank order correlation coefficient = 0.9134 at 0.01 level), the rankings do change significantly when faculty members in the middle third of the rankings are examined separately (Spearman rank order correlation coefficient = −0.45 at 0.01 level). In other words, Scopus significantly alters the relative ranking of those scholars that appear in the middle

[5]Table 3 also shows that WoS includes 391 quality of these citations (or 17.0%) more citations than Scopus (2,692 in comparison to 2,301, respectively), when citations from pre-1996 are counted.

TABLE 3. Impact of adding Scopus citations on faculty and school citation counts (1996–2005).

Research areas of individual faculty members[a]	WoS	Scopus	Union of WoS and Scopus	% Increase
Human-computer interaction	544	740	853	56.80
Citation analysis, informetrics, scholarly communication, and strategic intelligence	508	459	564	11.00
Computer-mediated communication, gender and information technology, and discourse analysis	273	313	365	33.70
E-commerce, information architecture, information policy and electronic networking	162	168	188	16.00
Bibliometrics, collection development and management, evaluation of library sources and services, and serials	123	108	137	11.40
Information seeking and use, design and impact of electronic information sources, and informetrics	122	111	128	4.90
Intelligent interfaces for information retrieval and filtering, knowledge discovery, and user modeling	118	129	154	30.50
Information visualization, data mining, and data modeling	115	133	165	43.50
Communities of practice	88	159	175	98.90
Classification and categorization, ontologies, metadata, and information architecture	83	80	93	12.00
Critical theory and documentation	35	37	42	20.00
Computational linguistics, computer-mediated communication, and sociolinguistics and language acquisition	32	38	44	37.50
Citation analysis, bibliometrics, and data retrieval and integration	29	21	31	6.90
Information retrieval and data integration	28	32	40	42.90
Information policy, social and organizational informatics, and research methods	28	31	34	21.40
Faculty members total	2,288	2,559	3,013	31.70
School total[b]	2,023	2,301	2,733	35.10

[a]Each row in the table represents a single faculty member and the main research topics covered by him or her. It would have been practically impossible to classify citations by individual topics rather than individual faculty members.

[b]Excludes duplicate citations.

TABLE 4. Impact of adding Scopus citations on the ranking of faculty members (1996–2005).

	WoS		Union of WoS and Scopus	
Faculty member	Count	Rank	Count	Rank
A	544	1	853	1
B	508	2	564	2
C	273	3	365	3
D	162	4	188	4
E	123	5	137	8
F	122	6	128	9
G	118	7	154	7
H	115	8	165	6
I	88	9	175	5
J	83	10	93	10
K	35	11	42	12
L	32	12	44	11
M	29	13	31	15
N	28	14T	40	13
O	28	14T	34	14

of the rankings but not for those at the top or bottom of the rankings.

- The overlap of SLIS citations between the two databases is relatively low—58.2% (see Figure 1) with significant differences from one research area to another ranging from a high 82.0% to a low 41.1% (see Table 5).
- The number of unique citations found in Scopus is noticeably high in comparison to that of WoS (710 or 26.0% in comparison to 432 or 15.8%, respectively) (see Figure 1). The overlap and uniqueness between the two databases is almost identical to what Whitley (2002) found in her study that compared the duplication (60%) and uniqueness of citing documents in Chemical Abstracts (23%) and Science Citation Index (17%).

Regarding the type of documents in which the citations were found, the main difference between the two databases is in the coverage of conference proceedings. Scopus retrieves considerably more citations from refereed conference papers than WoS (359 in comparison to 229, respectively; see Table 6). What is more important is that of all 496 citations from

DOI: 10.1002/asi

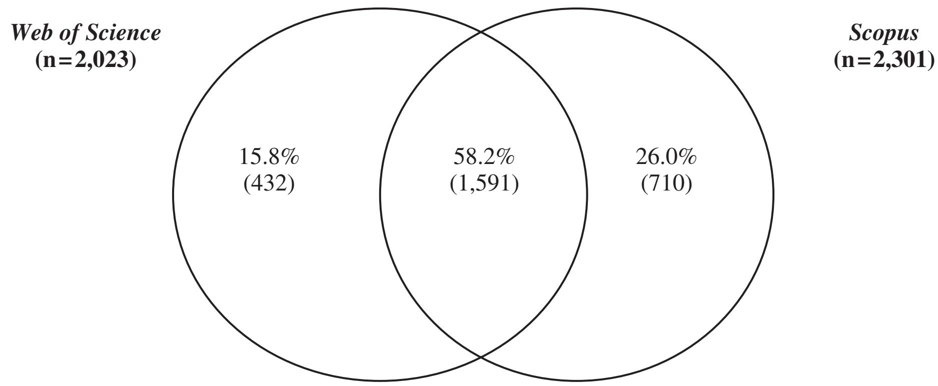

FIG. 1. Distribution of unique and overlapping citations in WoS and Scopus ($N = 2{,}733$).

TABLE 5. Overlap between Scopus and Web of Science (1996–2005).

Research areas of individual faculty members[a]	WoS	Scopus	Union	Overlap	%
Human-computer interaction	544	740	853	430	50.4
Citation analysis, informetrics, scholarly communication, and strategic intelligence	508	459	564	403	71.5
Computer-mediated communication, gender and information technology, and discourse analysis	273	313	365	221	60.5
E-commerce, information architecture, information policy and electronic networking	162	168	188	142	75.5
Bibliometrics, Collection development and management, evaluation of library sources and services, and serials	123	108	137	94	68.6
Information seeking and use, design and impact of electronic information sources, and informetrics	122	111	128	105	82.0
Intelligent interfaces for information retrieval and filtering, knowledge discovery, and user modeling	118	129	154	83	53.9
Information visualization, data mining, and data modeling	115	133	165	92	55.8
Communities of practice	88	159	175	72	41.1
Classification and categorization, ontologies, metadata, and information architecture	83	80	93	70	75.3
Critical theory and documentation	35	37	42	30	71.4
Computational linguistics, computer-mediated communication, and sociolinguistics and language acquisition	32	38	44	26	59.1
Citation analysis, bibliometrics, and data retrieval and integration	29	21	31	19	61.3
Information retrieval and data integration	28	32	40	20	50.0
Information policy, social and organizational informatics, and research methods	28	31	34	25	73.5
Faculty members total	2,288	2,559	3,013	1,832	60.8
School total[b]	2,023	2,301	2,733	1,591	58.2

[a]Each row in the table represents a single faculty member and the main research topics covered by him or her. It would have been practically impossible to classify citations by individual topics rather than individual faculty members.

[b]Excludes duplicate citations.

conference papers, 53.8% are uniquely found in Scopus in comparison to only 27.6% in WoS (19.6% of citations from conference papers are found in both databases). This can have significant implications for citation analysis and the evaluation of individual scholars, especially when those evaluated include authors who use conferences as a main channel of scholarly communication. Without Scopus, authors who communicate extensively through conferences will be at a disadvantage when their citation counts are compared with those who publish primarily in journals due to poor coverage of conference proceedings in WoS. Whether the value, weight, or quality of citations found in conference papers is

DOI: 10.1002/asi

TABLE 6. Web of Science and Scopus citation count by document type (1996–2005).

Document type	WoS Count[a]	WoS %	Scopus Count[a]	Scopus %	Union Count[a]	Union %
Journal articles	1,529	75.6	1,754	76.2	1,968	72.0
Conference papers	229	11.3	359	15.6	496	18.1
Review articles	172	8.5	147	6.4	175	6.4
Editorial materials	63	3.1	36	1.6	64	2.3
Book reviews	17	0.8	0	0.0	17	0.6
Letters to editors	9	0.4	2	0.1	9	0.3
Bibliographic essays	2	0.1	2	0.1	2	0.1
Biographical item	2	0.1	1	0.0	2	0.1
Total	2,023	100.0	2,301	100.0	2,733	100.0
Total from journals	1,794	88.7	1,942	84.4	2,237	81.9
Total from conference papers	229	11.3	359	15.6	496	18.1
Total	2,023	100.0	2,301	100.0	2,733	100.0

[a]Excludes duplicate citations.

TABLE 7. Google Scholar citation count by document type (1996–2005).[a]

Document type	Count[b]	%
Journal articles	2,215	40.32
Conference papers	1,849	33.66
Doctoral dissertations	261	4.75
Master's theses	243	4.42
Book chapters	199	3.62
Technical reports	129	2.35
Reports	110	2.00
Books	102	1.86
Review articles	86	1.57
Conference presentations	72	1.31
Unpublished papers	65	1.18
Bachelor's theses	34	0.62
Working papers	31	0.56
Editorial materials	25	0.46
Research reports	23	0.42
Workshop papers	15	0.27
Doctoral dissertation proposals	9	0.16
Conference posters	9	0.16
Book reviews	3	0.05
Master's thesis proposals	3	0.05
Preprints	3	0.05
Conference paper proposals	2	0.04
Government documents	2	0.04
Letters to the editor	2	0.04
Biographical item	1	0.02
Total	5,493	100.00
Total from journals	2,332	42.45
Total from conference papers	1,849	33.66
Total from journals and conference papers	4,181	76.12
Total from dissertations/theses	538	9.79
Total from books	301	5.48
Total from reports	262	4.77
Total from other document types	211	3.84
Total	5,493	100.00

[a]Table excludes 475 citations that did not have complete bibliographic information. These citations included bachelor theses, presentations, grant proposals, doctoral qualifying examinations, submitted manuscripts, syllabi, term papers, research proposals, working papers, Web documents, preprints, student portfolios, and so on.
[b]Excludes duplicate citations.

different from those of journal articles is not within the scope of this study; however, it should be emphasized that, as with journals, some conferences have stringent refereeing processes and low acceptance rates and others do not. All of the conference proceedings indexed by Scopus are peer-reviewed.

In conclusion, the findings suggest that to accurately evaluate and/or rank scholars, journals, and departments by way of citations, one should employ both WoS and Scopus to generate accurate citation accounts because these two databases largely complement rather than replace each other. Moreover, given the low overlap, or the high degree of uniqueness, in citations between the two databases, the findings further suggest that the use of Scopus in addition to WoS may have significant implications on journal impact factors (Garfield, 1996, 2006; Nisonger, 2004b) as well as on the correlation between citation data and perception-based evaluations or rankings (Meho & Sonnenwald, 2000; So, 1998). It will be important to verify the influence of Scopus data on journal impact factors in particular as they are often used by promotion and tenure committees, funding agencies, and collection development librarians for assessing the research impact/quality of scholars and journals (Moed, 2005).

Effect of Google Scholar on Citation Counts and Rankings of SLIS Faculty

Data collected in this study show that, in contrast to WoS and Scopus, which index citations mainly from journal articles and conference papers, citations found through GS come from many different types of documents, including journal articles, conference papers, doctoral dissertations, master's theses, technical reports, research reports, chapters, and books, among others (see Table 7). Data also show that the majority of citations found through GS come from documents published after 1993 (see Table 8). A main reason for this is that the study group has less citable works published before 1993 in comparison to those published since then. Another reason is that, unlike WoS and Scopus, which enter the citation information into their databases in a semi-automatic fashion, GS relies exclusively on the availability of online fulltext documents; therefore, retrospective coverage will increase only as older materials are converted into digital format and published on the Web. As mentioned earlier, analysis in this study is based only on citations found in journal items and conference papers published between 1996 and 2005.

Results show that GS identifies 1,448 (or 53.0%) more citations than WoS and Scopus combined (4,181 citations for GS in comparison to 2,733 for the union of WoS and Scopus). Results also show that combining citations from GS, WoS, and Scopus increases the number of citations to SLIS

DOI: 10.1002/asi

TABLE 8. Google Scholar citation distribution by year.

Years	Citations from journals and conference papers
1971–1975	1
1976–1980	1
1981–1985	9
1986–1990	29
1991	4
1992	12
1993	9
1994	43
1995	67
Subtotal	175
1996	101
1997	145
1998	176
1999	248
2000	350
2001	409
2002	539
2003	671
2004	752
2005	790
Subtotal	4,181
Total	4,356

faculty members as a whole by 93.4% (from 2,733 to 5,285 citations). In other words, one would miss over 93.4% of relevant citations if searching were limited to WoS and Scopus. Although the high number of unique citations in GS could be very helpful for those individuals seeking promotion, tenure, faculty positions, or research grants, most of these citations come from low-impact journals and/or conference proceedings (see below).

Data show that the percentage of increase in citation counts for SLIS faculty members varies considerably when GS results are added to those of WoS and Scopus (range = 120.2%). Faculty members with research strengths in the areas of communities of practice, computer-mediated communication, data mining, data modeling, discourse analysis, gender and information technology, human–computer interaction, information retrieval, information visualization, knowledge discovery, and user modeling had their citation counts increase considerably more than those faculty members with research strengths in other areas (see Table 9). Although this suggests that one should use GS to generate accurate citation counts of these authors, unlike the effect of adding Scopus' unique citations to those of WoS, adding GS's unique citations data to those of WoS and Scopus does not significantly alter the relative ranking of faculty members—Spearman rank order correlation coefficient = 0.976 at 0.001 level (see Table 10).

Even when GS results are added to those of WoS and Scopus separately, GS results do not significantly change the relative ranking of scholars—the Spearman rank order correlation coefficient between GS and WoS = 0.874 and between GS and Scopus = .970. Perhaps equally important is that the overlap between GS and the union of WoS and Scopus is considerably low (30.8%) and that GS misses a high number (1,104 or 40.4%) of the 2,733 citations found by WoS and Scopus (see Figure 2). Both of these figures are very striking, especially given the fact that virtually all citations from WoS and Scopus come from refereed and/or reputable sources.

To test that these results were not an outcome of the study group size and source, citation data was collected for 10 additional LIS faculty members, specializing in several research areas such as archives, children and young adult librarianship, digital libraries, evaluation of library services, health and medical information, public libraries, and school library media. These 10 faculty members were identified through searches in Library Literature and Information Science and WoS databases and were selected based on the number of refereed journal articles each one of them had in these databases (at least five articles in one or both databases). Results of this group of faculty members, who are cited in 442 documents (333 journal articles, 68 conference papers, 29 review articles, 8 editorial materials, 3 book reviews, and 1 letter), showed very similar patterns to those of the main study group (see Table 11). The only major difference found was that GS increases the citation count of the test group by 51.9% in comparison to 93.4% for the main study group. This difference is mostly attributable to the fact that some of the main sources of citations covering the research areas of the test group are either not yet available in

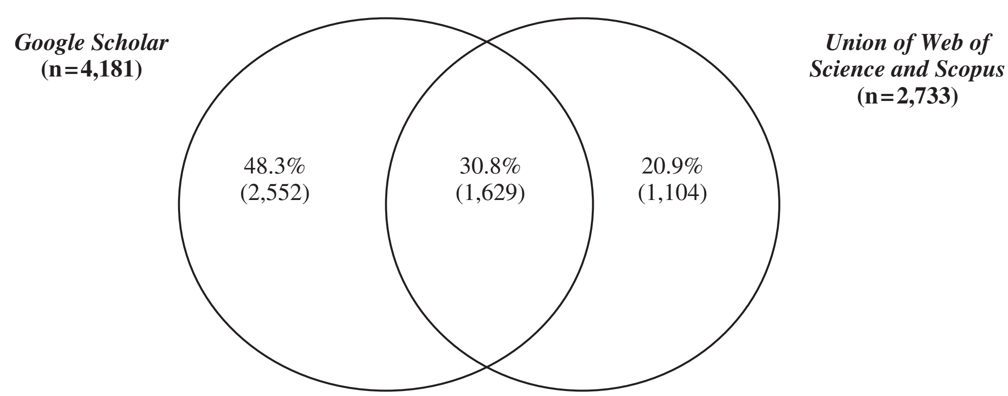

FIG. 2. Distribution of unique and overlapping citations in Google Scholar, Web of Science, and Scopus (N = 5,285).

DOI: 10.1002/asi

TABLE 9. Impact of adding Google Scholar citations on faculty members' citation count (1996–2005).

Research areas of individual faculty members[a]	Union of WoS and Scopus	GS	Union of three sources	% Increase
Human-computer interaction	853	1,786	2,078	143.6
Citation analysis, informetrics, scholarly communication, and strategic intelligence	564	517	802	42.2
Computer-mediated communication, gender and information technology, and discourse analysis	365	671	797	118.4
E-commerce, information architecture, information policy and electronic networking	188	164	244	29.8
Bibliometrics, collection development and management, evaluation of library sources and services, and serials	137	94	169	23.4
Information seeking and use, design and impact of electronic information sources, and informetrics	128	114	171	33.6
Intelligent interfaces for information retrieval and filtering, knowledge discovery, and user modeling	154	260	291	89.0
Information visualization, data mining, and data modeling	165	187	249	50.9
Communities of practice	175	342	403	130.3
Classification and categorization, ontologies, metadata, and information architecture	93	76	125	34.4
Critical theory and documentation	42	46	60	42.9
Computational linguistics, computer-mediated communication, and sociolinguistics and language acquisition	44	73	92	109.1
Citation analysis, bibliometrics, and data retrieval and integration	31	29	39	25.8
Information retrieval and data integration	40	46	59	47.5
Information policy, social and organizational informatics, and research methods	34	20	42	23.5
Faculty members total	3,013	4,425	5,621	86.6
School total[b]	2,733	4,181	5,285	93.4

Note. Included in this table are citations from journals and conference papers only. Excluded are citations from dissertations, theses, reports, books, conference presentations, meeting abstracts, research and technical reports, unpublished papers, working papers, workshop papers, and so on. The table also excludes 475 citations that did not have complete bibliographic information. These citations included: bachelor theses, presentations, grant proposals, doctoral qualifying examinations, submitted manuscripts, syllabi, term papers, research proposals, working papers, Web documents, preprints, student portfolios, and so on.

[a]Each row in the table represents a single faculty member and the main research topics covered by him/her. It would have been practically impossible to classify citations by individual topics rather than individual faculty members.

[b]Excludes duplicate citations.

digital format (e.g., *The American Archivist* and *Archivaria*) or have only the last few years of their issues available online (e.g., *Archival Issues*, 1997–present; *Archival Science*, 2000–present; and *Journal of Education for Library and Information Science*, 2003–present). This provides evidence that GS favors those faculty members who work in research areas that have high Web presence or those who make their scholarly works available online more than those faculty members who work in research areas that do not have high Web presence or who do not post their works online.

Given the results from both the main study group and the test group, as well as the fact that GS is so cumbersome to use, misses a significant number of citations from refereed sources, and has little or no influence on the relative rankings of scholars, one could conclude that, as far as the two study groups are concerned, GS is superfluous to using both WoS and Scopus to generate citation counts and rankings for assessing and comparing scholars, journals, and academic departments, especially when the focus of a study is on citations in refereed journals and conference proceedings. Results of this study also show that the use of Scopus in addition to WoS further diminishes the value of using GS as evidenced in the increase in Spearman's rho correlation from .874 between GS and WoS to 0.976 between GS and the union of WoS and Scopus.

Considering the type of documents in which the citations were found, GS retrieves significantly more (almost 4 times as many) citations from conference papers than WoS and Scopus combined (1,849 in comparison to 496, respectively). In contrast, WoS and Scopus retrieve almost as many citations from journals as GS does (1,968 in comparison to 2,332, respectively). It should be emphasized here that the

DOI: 10.1002/asi

TABLE 10. Impact of adding Google Scholar citations on the ranking of faculty members (1996–2005).

	Union of WoS and Scopus		Union of WoS, Scopus, and GS	
Faculty member	Count	Rank	Count	Rank
A	853	1	2,078	1
B	564	2	802	2
C	365	3	797	3
D	188	4	244	7
I	175	5	402	4
H	165	6	249	6
G	154	7	291	5
E	137	8	169	9
F	128	9	171	8
J	93	10	125	10
L	44	11	92	11
K	42	12	60	12
N	40	13	59	13
O	34	14	42	14
M	31	15	39	15

Note. WoS = Web of Science; GS = Google Scholar.

TABLE 11. Comparison of results between main study group and test group (1996–2005).

	Main study group	Test group
Ratio of citations between WoS and Scopus	.88 : 1.0	.91 : 1.0
Percent of increase when adding Scopus unique citations to those of WoS	35.1	40.6
Percent of overlap between WoS and Scopus	58.2	50.5
Percent of unique citations found in Scopus in the union of WoS and Scopus	26.0	27.8
Percent of unique citations found in WoS in the union of WoS and Scopus	15.8	21.6
Percent of increase when adding GS unique citations to those of WoS and Scopus	93.4	51.9
Percent of overlap between GS and union of WoS and Scopus	30.8	26.2
Percent of WoS and Scopus citations missed by GS	40.4	60.8
Spearman Rank Order correlation coefficient between GS and union of WoS and Scopus	.976	.967

Note. WoS = Web of Science; GS = Google Scholar. Spearman rank order correlation coefficient between GS and union of WoS and Scopus for all 25 faculty members (i.e., main study group and test group combined) = .982.

relatively poor coverage of conference papers by WoS and Scopus, or the relative good coverage of this document type by GS, has much to do with the fact that many authors make their conference papers available online themselves. Almost half of GS's unique citations from conference papers and many of its citations from journals were identified through fulltext documents made available online by their authors (i.e., self-archived) rather than from the official Web sites of the publishers of the conference proceedings and journals.[6] What these kinds of findings reveal is that there is a dramatic advantage in favor of the articles that their authors make available online. According to Harnad and Brody (2004):

> The proportion of articles for which their authors provide OA [Open Access] is likely to increase dramatically now, in part because of the mounting evidence for the impact advantage OA confers. OA will also increase because of the growing number of journals that have already given their official "green light" to author self-archiving, partly because journal impact factors also benefit from increased article impact, and partly because journals are eager to demonstrate that they have no wish to stand in the way of OA and its benefits to research and researchers.

Sources of Citations and Their Refereed Status and Language

As mentioned earlier, only 58.2% (or 1,591) of all WoS and Scopus citations (n = 2,733) were duplicated in both databases, raising the important question of where the 1,142 unique citations originated from. Answering this question will identify important LIS-relevant journals and conference proceedings that are not indexed by the databases. This list could be very useful for collection development librarians as well as for the producers of the databases should they decide

[6]In these cases, data collection involved going to the root Web site to identify the full bibliographic information of the citing documents. Most often, the root Web sites were the curricula vitae of the authors of the citing documents.

DOI: 10.1002/asi

to increase their coverage of LIS literature. Data show that the 2,023 citations from WoS come from 505 different journals and conference proceedings, whereas the 2,301 citations from Scopus come from 681 different titles. The 2,733 citations from the union of both databases come from 787 different journals and conference proceedings. Of these 787 titles, 398 (or 50.5%) are indexed by both databases, 107 (or 13.5%) are indexed only by WoS, and 283 (or 36.0%) are indexed only by Scopus.

Data show that the top 54 (or 6.9%) sources of citations in WoS and/or Scopus account for 1,410 (or 51.6%) of all of the databases' 2,733 citations for the study group, reflecting the Matthew Effect in citations—a small number of sources attracts the lion's share of citations and a large number of sources receives relatively few (Merton, 1968). Of these 54 sources of citations, 10 (or 18.5%) are not indexed by WoS, whereas only one is not indexed by Scopus. Of the 10 titles not indexed by WoS, six are conference proceedings and four are journals (see Table 12). It is interesting to note that these four journals have higher impact factor scores than most of the LIS journals that are currently indexed by WoS and/or included in ISI's Journal Citation Reports—*Computer Supported Cooperative Work* (1.190), *Internet and Higher Education* (1.022), *D-Lib Magazine* (.615), and *International Review of Research in Open and Distance Learning* (.354), in comparison to a median impact factor of 0.480 for LIS journals (Thomson Scientific, 2006).[7] As for the six conference proceedings, three were from the ACM and the IEEE.

These results, which are influenced by the makeup of our study group (i.e., one with strong research focus in communities of practice, computer-mediated communication, human–computer interaction, and information visualization, in addition to traditional LIS research areas), suggest that if WoS is to reduce the gap in its coverage of LIS and LIS-related fields, it should consider adding at least the relevant high-impact factor journals and conference proceedings that Scopus uniquely indexes. As is, the results imply that Scopus is necessary to use along with WoS for providing a better and more accurate picture of the impact SLIS research makes on other fields, as evidenced by the computer science, education, and engineering titles that cite SLIS literature and are only/primarily indexed by Scopus. Although SLIS does not constitute a representative sample from which to generalize the findings and conclusions to the entire LIS field, similar results and conclusions may be found or made when citations to the works of faculty members in other LIS schools are examined.

Further analysis shows that when a journal or a conference proceeding is indexed by both WoS and Scopus, the former tends to identify more citations from these commonly indexed sources than the latter does in the majority of cases. For example, WoS identifies 145 citations from the *Journal of the American Society for Information Science and Technology*, whereas Scopus finds only 112 from the same journal covering the same time period (see Table 12 for more examples). There are, however, cases where Scopus identifies more citations than WoS from the same titles (e.g., *Journal of Computer-Mediated Communication—JCMC*, *Journal of Educational Computing Research*, and *Education for Information*). Reasons for these variations in coverage between the two databases include database errors (e.g., lack of cited references information and incomplete lists of references in some database records), partial indexing of journal content (e.g., not indexing all articles published in a journal and not indexing book reviews as is the case in Scopus, although some of these items contain citations), and incomplete coverage period of journals (e.g., missing an entire issue or volume—*JCMC*, for example is covered by Scopus from 1996 to present, whereas it was just recently added to WoS covering the years 2005 to present).

As in the case of Scopus, GS results also raise the important question of where it found the 2,552 citations that were missed by both WoS and Scopus. As mentioned earlier, GS is able to search documents from hundreds of publishers, including items their authors themselves have made available online. An examination of the top 51 sources of citations found exclusively in GS, however, shows that 14 are actually indexed by Scopus and six are indexed by both Scopus and WoS (see Table 13). The reasons why WoS and Scopus miss some citations from these 20 titles are similar to those mentioned above (e.g., database errors, partial and incomplete coverage, etc.).

Results also show that 10 of the remaining top 51 sources of GS unique citations are journals and 21 are conference proceedings. To identify the quality or impact of these 31 titles, we used Scopus to generate citation counts to these titles and compared the counts to those of highly ranked LIS journals and conference proceedings, such as the *Annual Review of Information Science and Technology*, *College & Research Libraries*, *Information Processing & Management*, *Journal of Documentation*, *Journal of the American Society for Information Science and Technology*, and *Scientometrics* (Nisonger & Davis, 2005; Thomson Scientific, 2006). Results show that with the exception of one title, none of these 31 titles is cited more than the top 11 LIS journals and conference proceedings as identified by Nisonger and Davis (see Table 14). This finding raises important questions regarding the quality of citations uniquely found in GS, as well as the wisdom of using these citations for tenure and promotion purposes, despite the fact that most of the citations uniquely found by GS are from refereed sources (only 2 of the top 51 sources are not refereed). Note that of the top 51 sources of citations in GS, 15 are published or sponsored by the ACM, three by the IEEE, and three jointly by the ACM and IEEE (see Table 13).

Another important finding is that GS provides significantly better coverage of non-English language materials (6.94% of its total citations) than both WoS (1.14%) and Scopus (0.70%; see Table 15). This discovery suggests that

[7]The 2006 citation impact factor formula = number of citations in 2006 to articles published in 2005 + number of citations in 2006 to articles published in 2004, divided by number of articles published in 2004–2005.

DOI: 10.1002/asi

TABLE 12. Sources of citations in Scopus and Web of Science (1996–2005).[a]

Title	WoS	Scopus	Union	Rank	% (n = 2,733)
Journal of the American Society for Information Science and Technology	145	112	147	1	5.4
Lecture Notes in Computer Science series[a]	118	33	118	2	4.3
Scientometrics	78	69	79	3	2.9
Journal of Documentation	71	56	71	4	2.6
Annual Review of Information Science and Technology	53	51	53	5	1.9
Journal of Information Science	46	47	47	6	1.7
Proceedings of the Annual Meeting of the American Society for Information Science and Technology	41	35	45	7	1.6
Journal of Computer-Mediated Communication	15	41	41	8	1.5
International Journal of Human–Computer Studies	40	40	40	9	1.5
Lecture Notes in Artificial Intelligence Series[b]	40	7	40	10	1.5
Information Processing & Management	39	38	39	11	1.4
Interacting With Computers	34	34	34	12	1.2
Library & Information Science Research	31	27	32	13	1.2
Information Society: An International Journal	24	25	28	14	1.0
Aslib Proceedings	22	14	23	15	0.8
Behaviour and Information Technology	22	23	23	16	0.8
Computers & Education	22	21	22	17	0.8
College & Research Libraries	21	17	21	18	0.8
Library Trends	21	21	21	19	0.8
Computers in Human Behavior	20	20	20	20	0.7
Information Research: An International Electronic Journal	16	19	20	21	0.7
Journal of Academic Librarianship	19	18	20	22	0.7
Cyberpsychology & Behavior	18	18	19	23	0.7
Internet Research	19	16	19	24	0.7
Computer Supported Cooperative Work: The Journal of Collaborative Computing	Not indexed	18	18	25	0.7
Internet and Higher Education	Not indexed	18	18	26	0.7
Knowledge Organization	18	15	18	27	0.7
ACM Conference on Computer Supported Cooperative Work (CSCW)	Not indexed	16	16	28	0.6
Journal of Educational Computing Research	4	16	16	29	0.6
Library Collections, Acquisitions, and Technical Services	16	10	16	30	0.6
Human-Computer Interaction	15	14	15	31	0.5
SIGCHI Conference on Human Factors in Computing Systems (ACM)	Not indexed	15	15	32	0.5
Journal of Computer Assisted Learning	14	14	14	33	0.5
Libri	14	Not indexed	14	34	0.5
SPIE Proceedings Series (International Society for Optical Engineering)	Not indexed	14	14	35	0.5
Education for Information	2	11	13	36	0.5
ETR&D—Educational Technology Research & Development	12	11	13	37	0.5
Library Quarterly	13	12	13	38	0.5
Proceedings of the National Academy of Sciences	13	13	13	39	0.5
Government Information Quarterly	11	10	12	40	0.4
Library Resources & Technical Services	12	10	12	41	0.4
Online Information Review	12	11	12	42	0.4
American Society for Engineering Education (ASEE) Annual Conference	Not indexed	11	11	43	0.4
Canadian Journal of Information and Library Science	10	11	11	44	0.4
D-Lib Magazine	Not indexed	11	11	45	0.4
International Review of Research in Open and Distance Learning	Not indexed	11	11	46	0.4
Language Learning & Technology	9	11	11	47	0.4
Serials Librarian	9	10	11	48	0.4
ACM/IEEE Joint Conference on Digital Libraries / ACM Conference on Digital Libraries	Not indexed	10	10	49	0.4
Annual Meeting of the Human Factors and Ergonomics Society	Not indexed	10	10	50	0.4
Bioinformatics	9	10	10	51	0.4
Educational Technology & Society	5	10	10	52	0.4
International Journal of Information Management	10	8	10	53	0.4
Journal of the American Medical Informatics Association	10	9	10	54	0.4
Total	1,193	1,152	1,410		51.6

Note. WoS = Web of Science; GS = Google Scholar.

[a] Highly cited titles discovered through the test group, but not through the main study group include *School Library Media Research* (14), *Medical Reference Services Quarterly* (9), *Archival Science* (8), and *OCLC Systems & Services* (6). All four are indexed by, or found through, Scopus and GS, but not WoS.

[b] Citations published in *LNCS* and *LNAI* series come from a number of different conference proceedings, such as Advances in Case-Based Reasoning: European Workshop, CHI conferences, International Conference on Case-Based Reasoning Research and Development, European Conference on Research and Advanced Technology for Digital Libraries, and International Conference on Web-Based Learning.

Title changes of journals and conference proceedings were taken into consideration in the calculation of citation count.

DOI: 10.1002/asi

TABLE 13. Sources of citations unique to Google Scholar (1996–2005).

Source	Status	Count
Annual Hawaii International Conference on System Sciences (HICSS) (IEEE)	**Refereed**	**57**
Ciencia da Informacao	Refereed	35
CHI Extended Abstracts on Human Factors in Computing Systems (ACM)	Refereed	30
SIGCHI Conference on Human Factors in Computing Systems (ACM)	**Refereed**	**30**
HCI International: International Conference on Human–Computer Interaction	Refereed	27
Conference on Interaction Design and Children (ACM)	Refereed	23
OZCHI: Australian Conference on Computer–Human Interaction (IEEE)	Refereed	23
Lecture Notes in Computer Science (ACM, IEEE, and others)	**Refereed**	**22**
Annual Conference of the Australasian Society for Computers in Learning in Tertiary Education – ASCILITE	Refereed	21
Interact: IFIP TC13 International Conference on Human Computer Interaction	Refereed	20
European Conference on Computer Supported Cooperative Work	Refereed	19
Information Systems Research Seminar in Scandinavia (IRIS)	Refereed	18
ACM Conference on Computer Supported Cooperative Work	**Refereed**	**17**
Australasian Journal of Educational Technology / Australian Journal of Educational Technology	Refereed	15
International ACM SIGGROUP Conference on Supporting Group Work (Group)	**Refereed**	**14**
Internet Research Annual: Association of Internet Researchers (AOIR) Annual Conference	Refereed	14
Nordic Conference on Human–Computer Interaction	Refereed	14
Personal and Ubiquitous Computing (ACM)	**Refereed**	**14**
International Conference on Human Computer Interaction with Mobile Devices and Services (Mobile HCI) (ACM)	Refereed	13
Proceedings of the Annual Meeting of the American Society for Information Science and Technology	**Refereed**	**13**
ACM/IEEE Joint Conference on Digital Libraries / ACM Conference on Digital Libraries	**Refereed**	**12**
Computer Assisted Language Learning	Refereed	11
Americas Conference on Information Systems (AIS)	Refereed	10
Education and Information Technologies	Refereed	10
IFLA General Conference	Not refereed	10
Text Retrieval Conference (TREC)	Not refereed	10
ACM SIGGROUP Bulletin	Refereed	9
IEEE International Conference on Information Visualisation	**Refereed**	**9**
Interactions (ACM)	Refereed	9
International Conference on Artificial Intelligence in Education (AIED)	Refereed	9
Journal of the American Society for Information Science and Technology	**Refereed**	**9**
Annual Meeting of the American Educational Research Association (AERA)	Refereed	8
Cognition	**Refereed**	**8**
Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques (DIS) (ACM)	Refereed	8
Distance Education: An International Journal	**Refereed**	**8**
International Conference on Ubiquitous Computing (UbiComp)	**Refereed**	**8**
ACM International Conference of Computer-Supported Collaborative Learning (CSCL)	Refereed	7
ACM/IEEE International Conference on Software Engineering (ICSE)	**Refereed**	**7**
Educational Media International	Refereed	7
Ethics and Information Technology	Refereed	7
ACM SIGCHI Bulletin	Refereed	6
ACM Transactions on Computer–Human Interaction	**Refereed**	**6**
American Journal of Distance Education	Refereed	6
Annual International ACM SIGIR Conference on Research and Development in Information Retrieval	**Refereed**	**6**
Collection Building	**Refereed**	**6**
Ed-Media: World Conference on Educational Multimedia	Refereed	6
Interacción	**Refereed**	**6**
Issues in Information Systems	**Refereed**	**6**
Journal of Educational Technology Systems	**Refereed**	**6**
Latin American Conference on Human–Computer Interaction	Refereed	6
Perspectivas em Ciência da Informação	Refereed	6

Note. Titles in bold are journals/conference proceedings indexed by Web of Science and/or Scopus. Title changes of journals and conference proceedings were taken into consideration in the calculation of citation count.

DOI: 10.1002/asi

TABLE 14. Comparison between citation counts of top 46 sources of Google Scholar unique citations and those of top library and information science journals and conference proceedings.

Title	Citation count	Rank
Journal of the American Society for Information Science and Technology	6,679	1
Information Processing and Management	5,164	2
Annual Meeting of the American Educational Research Association (AERA)	**4,357**	**3**
Journal of Information Science	2,917	4
Scientometrics	2,596	5
Journal of Documentation	2,507	6
Proceedings of the Annual Meeting of the American Society for Information Science and Technology	1,654	7
Journal of Academic Librarianship	1,459	8
Annual Review of Information Science and Technology	1,349	9
Library Quarterly	1,268	10
College & Research Libraries	1,197	11
Library & Information Science Research	1,023	12
European Conference on Computer Supported Cooperative Work	**872**	**13**
Ed–Media: World Conference on Educational Multimedia	**866**	**14**
International Conference on Ubiquitous Computing (UbiComp)	**824**	**15**
HCI International: International Conference on Human–Computer Interaction	**783**	**16**
International Conference of Computer–Supported Collaborative Learning (CSCL) (ACM)	**703**	**17**
American Journal of Distance Education	**684**	**18**
Americas Conference on Information Systems (AIS)	**637**	**19**

Note. Source of citation count data for this table: Scopus (1996–2006). All of the nonbold items are either fully or selectively indexed by Web of Science and/or ScopuS and are top-ranked library and information science journals and conference proceedings, according to *Journal Citation Reports* (Thomson Scientific, 2006) and Nisonger and Davis (2005). All of the bold items are sources uniquely searched/identified by Google Scholar.

Title changes of journals and conference proceedings were taken into consideration in the calculation of citation count.

TABLE 15. Citation count distribution by language (1996–2005).

	WoS		Scopus		GS		Total	
	Count	%	Count	%	Count*	%	Count*	%
English	2,000	98.86	2,285	99.30	3,891	93.06	4,972	94.08
Portuguese					92	2.20	92	1.74
Spanish	4	0.20	3	0.13	63	1.51	68	1.29
German	13	0.64	9	0.39	38	0.91	50	0.95
Chinese					44	1.05	44	0.83
French	3	0.15	1	0.04	32	0.77	35	0.66
Italian					8	0.19	8	0.15
Japanese	3	0.15	3	0.13	1	0.02	4	0.08
Swedish					3	0.07	3	0.06
Czech					2	0.05	2	0.04
Dutch					2	0.05	2	0.04
Finnish					2	0.05	2	0.04
Croatian					1	0.02	1	0.02
Hungarian					1	0.02	1	0.02
Polish					1	0.02	1	0.02
Total	2,023	100.00	2,301	100.00	4,181	100.00	5,285	100.00
Non-English	23	1.14	16	0.70	290	6.94	313	5.92

Note. WoS = Web of Science; GS = Google Scholar.

GS is indispensable for showing one's international impact, at least as far as SLIS faculty members are concerned.

CONCLUSIONS AND IMPLICATIONS

This study provides several useful suggestions for scholars conducting citation analysis and those who need assistance in compiling their own citation records. It informs researchers, administrators, editors, reviewers, funding agencies, and information professionals of the wisdom and/or need and value of using multiple sources for, and ways of, identifying citations to authors, articles, and journals. The study found that the addition of Scopus citations to those of WoS could significantly alter the ranking of authors. The study also found that GS stands out in its coverage of conference proceedings as well as international, non-English language journals, among others. Google Scholar also indexes a wide variety of document types, some of which may be of significant value to researchers and others. The use of Scopus and GS, in addition to WoS, reveals a more comprehensive and accurate picture of the extent of the scholarly relationship between SLIS (and by extension LIS) and other fields, as evidenced by the unique titles that cite SLIS literature (e.g., titles from the fields of Cognitive Science, Computer Science, Education, and Engineering, to name a few). Significantly, this study has demonstrated that:

1. Although WoS remains an indispensable citation database, it may be necessary to additionally use Scopus for locating citations to an author or title, and, by extension, journals, departments, and countries; as far as SLIS faculty members are concerned, the use of Scopus as an additional citation source significantly influences their citation counts and rankings.
2. Although Scopus provides more comprehensive citation coverage of LIS and LIS-related literature than WoS for the period 1996–2005, the two databases complement rather than replace each other.
3. Although both Scopus and GS help identify a considerable number of citations not found in WoS, only Scopus significantly alters the ranking of SLIS authors as measured by WoS.
4. Although GS unique citations are not of the same quality or weight as those found in WoS or Scopus, they could be very useful in showing evidence of broader international impact than could possibly be done through the two proprietary databases.
5. The GS value for citation searching purposes is severely diminished by its inherent problems. The GS data are not limited to refereed, high-impact journals and conference proceedings. Google Scholar is also very cumbersome to use and needs significant improvement in the way it displays search results and the downloading capabilities it offers for it to become a useful tool for large-scale citation analyses.
6. Given the low overlap or high uniqueness between the three tools, they may all be necessary to develop more accurate maps or visualizations of scholarly networks and impact both within and between research groups, journals, disciplines, and other research entities (Börner, Chen, & Boyack, 2003; Börner, Sanyal, & Vespignani, 2006; Small, 1999; White & McCain, 1997).
7. Each database or tool requires specific search strategy(ies) to collect citation data, some more accurately and quickly (i.e., WoS and Scopus) than others (i.e., GS).

This study has significant implications for funding agencies as well as editors and publishers of journals who may wish to use citation counts and rankings to identify subject experts to review grant applications or submitted manuscripts and to determine the impact of projects and articles they funded or published. The study has also significant implications for the wider scholarly community as researchers begin to adopt the methods and databases described or listed here to identify citations that may otherwise remain unknown. Continuous advances in information technology and improvements in online access to citation data suggest that future studies should explore:

- Databases and tools that can be used to locate citations from refereed sources not covered by WoS or Scopus
- The potential impact of these databases and tools on citation counts and rankings as well as the correlation between citation counts/rankings and peer reviews/assessments of publication venues
- Whether broader sourcing of citations alters an article, author, journal, or a department's relative ranking vis-à-vis others and, if so, how
- Which sources of citations provide better coverage of certain research areas than others
- The intrinsic quality, weight, or value of citations found in these sources

The recent emergence of Scopus and other citation databases and tools marks the beginning of a new era in citation analysis, an era that should help provide better services from the producers of these databases as they will compete for market share. Such competition will force database producers, such as Elsevier and Thomson Scientific, to constantly monitor and/or broaden their coverage of high impact journals and conference proceedings.

Acknowledgments

This project was partially funded by OCLC Library and Information Science Research Grant Project.

We are most grateful to Shahriar Akram, Kathleen Burlingame, Sara Franks, Lana Gottschalk, Sarah Horowitz, Mary Snyder, and Jacqueline Steinhofer for assistance with data collection and to Blaise Cronin, Alice Robbin, Debora Shaw, and three anonymous reviewers for helpful comments on the paper.

References

Adkins, D., & Budd, J. (2006). Scholarly productivity of U.S. LIS faculty. Library & Information Science Research, 28(3), 374–389.

Aksnes, D.W., & Taxt, R.E. (2004). Peer reviews and bibliometric indicators: A comparative study at a Norwegian university. Research Evaluation, 13(1), 33–41.

Association for Library and Information Science Education. (2006). LIS research areas classification scheme. Retrieved March 15, 2006, from http://www.alise.org/research_class_guide.html

Bakkalbasi, N., Bauer, K., Glover, J., & Wang, L. (2006). Three options for citation tracking: Google Scholar, Scopus and Web of Science. Biomedical Digital Libraries, 7. Retrieved October 04, 2006, from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1533854

Ballard, S., & Henry, M. (2006). Citation searching: New players, new tools. The Searcher: the Magazine for Database Professionals, 14(9), 24–33.

Bar-Ilan, J. (2006). H-index for Price medalists revisited. ISSI Newsletter, 2(1), 3–4.

Bauer, K., & Bakkalbasi, N. (2005). An examination of citation counts in a new scholarly communication environment. D-Lib Magazine, 11(9). Retrieved March 25, 2006, from http://www.dlib.org/dlib/september05/bauer/09bauer.html

Borgman, C.L., & Furner, J. (2002). Scholarly communication and bibliometrics. Annual Review of Information Science and Technology, 36, 3–72.

DOI: 10.1002/asi

Börner, K., Chen, C., & Boyack, K.W. (2003). Visualizing knowledge domains. Annual Review of Information Science and Technology, 37, 179–255.

Börner, K., Sanyal, S., & Vespignani, A. (2006). Network science: A theoretical and practical framework. Annual Review of Information Science and Technology, 40.

Bornmann, L., & Daniel, H-D. (2007). What do we know about the h-index? Journal of the American Society for Information Science and Technology, 58(9), 1381–1385.

Bornmann, L., & Daniel, H-D. (2005). Does the h-index for ranking of scientists really work? Scientometrics, 65(3), 391–392.

Budd, J.M. (2000). Scholarly productivity of U.S. LIS faculty: An update. The Library Quarterly, 70(2), 230–245.

Cronin, B. (1984). The citation process: The role and significance of citations in scientific communication. London: Taylor Graham.

Cronin, B., & Meho, L. (2006). Using the h-index to rank influential information scientists. Journal of the American Society for Information Science and Technology, 57(9), 1275–1278.

Cronin, B., Snyder, H., & Atkins, H. (1997). Comparative citation rankings of authors in monographic and journal literature: A study of sociology. Journal of Documentation, 53(3), 263–273.

de Arenas, J.L., Castanos-Lomnitz, H., & Arenas-Licea, J. (2002). Significant Mexican research in the health sciences: A bibliometric analysis. Scientometrics, 53(1), 39–48.

Elsevier Science Publishers. (2006). Scopus [Electronic database]. Amsterdam/New York: Author.

Funkhouser, E.T. (1996). The evaluative use of citation analysis for communications journals. Human Communication Research, 22(4), 563–574.

Gardner, S., & Eng, S. (2005). Gaga over Google? Scholar in the social sciences. Library Hi Tech News, 22(8), 42–45.

Garfield, E. (2006). The history and meaning of the journal impact factor. Journal of the American Medical Association, 295(1), 90–93.

Garfield, E. (1996). How can impact factors be improved? British Medical Journal, 313(7054), 411–413.

Glänzel, W. (1996). The needs for standards in bibliometric research and technology. Scientometrics, 35(2), 167–176.

Goodman, D., & Deis, L. (2005). Web of Science (2004 version) and Scopus. The Charleston Advisor, 6(3). Retrieved March 25, 2006, from http://www.charlestonco.com/dnloads/v6n3.pdf

Goodrum, A.A., McCain, K.W., Lawrence, S., & Giles, C.L. (2001). Scholarly publishing in the Internet age: A citation analysis of computer science literature. Information Processing & Management, 37(5), 661–675.

Harnad, S., & Brody, T. (2004). Comparing the impact of open access (OA) vs. non-OA articles in the same journals. D-Lib Magazine, 10(6). Retrieved on April 27, 2006, from http://www.dlib.org/dlib/june04/harnad/06harnad.html

Hirsch, J.E. (2005). An index to quantify an individual's scientific research output. Proceedings of the National Academy of Sciences of the United States of America, 102(46), 16569–16572. Retrieved February 15, 2006, from http://www.pnas.org/cgi/reprint/102/46/16569.

Holden, G., Rosenberg, G., & Barker, K. (2005). Bibliometrics: A potential decision making aid in hiring, reappointment, tenure and promotion decisions. Social Work in Health Care, 41(3–4), 67–92.

Jacsó, P. (2005a). As we may search—comparison of major features of the Web of Science, Scopus, and Google Scholar citation-based and citation-enhanced databases. Current Science, 89(9), 1537–1547. Retrieved March 15, 2006, from http://www.ias.ac.in/currsci/nov102005/1537.pdf.

Jacsó, P. (2005b). Google Scholar: The pros and the cons. Online Information Review, 29(2), 208–214.

Jacsó, P. (2006). Deflated, inflated and phantom citation counts. Online Information Review, 30(3), 297–309.

Kostoff, R.N. (1996). Performance measures for government-sponsored research: Overview and background. Scientometrics, 36(3), 281–292.

MacRoberts, M.H., & MacRoberts, B.R. (1996). Problems of citation analysis. Scientometrics, 36(3), 435–444.

Martin, B.R. (1996). The use of multiple indicators in the assessment of basic research. Scientometrics, 36(3), 343–362.

Meho, L.I., & Sonnenwald, D.H. (2000). Citation ranking versus peer evaluation of senior faculty research performance: A case study of Kurdish scholarship. Journal of the American Society for Information Science, 51(2), 123–138.

Merton, R.K. (1968). The Matthew Effect in science. Science, 159(3810), 56–63.

Moed, H.F. (2005). Citation analysis in research evaluation. Berlin/Heidelberg/New York: Springer.

Narin, F. (1976). Evaluative bibliometrics: The use of publication and citation analysis in the evaluation of scientific activity. Cherry Hill, NJ: Computer Horizons.

Nisonger, T.E. (2004a). Citation autobiography: An investigation of ISI database coverage in determining author citedness. College & Research Libraries, 65(2), 152–163.

Nisonger, T.E. (2004b). The benefits and drawbacks of impact factor for journal collection management in libraries. The Serials Librarian, 47(1–2), 57–75.

Nisonger, T.E., & Davis, C.H. (2005). The perception of library and information science journals by LIS education deans and ARL library directors: A replication of the Kohl–Davis study. College & Research Libraries, 66(4), 341–377.

Noruzi, A. (2005). Google Scholar: The new generation of citation indexes. Libri, 55(4), 170–180.

Oppenheim, C. (2007). Using the h-index to rank influential British researchers in information science and librarianship. Journal of the American Society for Information Science and Technology, 58(2), 297–301.

Pauly, D., & Stergiou, K.I. (2005). Equivalence of results from two citation analyses: Thomson ISI's citation index and Google's scholar service. Ethics in Science and Environmental Politics, 33–35.

Persson, O., & Åström, F. (2005). Most cited universities and authors in library & information Science 1990–2004. Bibliometric Notes, 7(2). Retrieved November 15, 2006, from http://www.umu.se/inforsk/BibliometricNotes/BN2-2005/BN2-2005.htm

Reed, K.L. (1995). Citation analysis of faculty publications: Beyond Science Citation Index and Social Science [sic] Citation Index. Bulletin of the Medical Library Association, 83(4), 503–508.

Roth, D.L. (2005). The emergence of competitors to the Science Citation Index and Web of Science. Current Science, 89(9), 1531–1536. Retrieved March 15, 2006, from http://www.ias.ac.in/currsci/nov102005/1531.pdf

Saha, S., Saint, S., & Christakis, D.A. (2003). Impact factor: A valid measure of journal quality? Journal of the Medical Library Association, 91(1), 42–46.

Seglen, P.O. (1998). Citation rates and journal impact factors are not suitable for evaluation of research. Acta Orthopaedica Scandinavica, 69(3), 224–229.

Small, H. (1999). Visualizing science by citation mapping. Journal of the American Society for Information Science, 50(9), 799–813.

So, C.Y.K. (1998). Citation ranking versus expert judgment in evaluating communication scholars: Effects of research specialty size and individual prominence. Scientometrics, 41(3), 325–333.

Thomson Corporation. (2006a). Web of Science 7.0. Retrieved June 15, 2005, from http://scientific.thomson.com/support/products/wos7/

Thomson Corporation. (2006b). University science indicators, 1981–2005. In-cites. Retrieved May 10, 2006, from http://in-cites.com/research/2006/may_8_2006-1.html

Thomson Scientific. (2006). Journal citation reports [Electronic database]. Philadelphia: Author.

van Raan, A.F.J. (1996). Advanced bibliometric methods as quantitative core of peer-review based evaluation and foresight exercises. Scientometrics, 36(3), 397–420.

van Raan, A.F.J. (2000). The Pandora's box of citation analysis: Measuring scientific excellence—the last evil? In B. Cronin & H.B. Atkins (Eds.). The web of knowledge: A festschrift in honor of Eugene Garfield (pp. 301–319). Medford, NJ: Information Today.

van Raan, A.F.J. (2005). Fatal attraction: Conceptual and methodological problems in the ranking of universities by bibliometric methods. Scientometrics, 62(1), 133–143.

Wallin, J.A. (2005). Bibliometric methods: Pitfalls and possibilities. Basic & Clinical Pharmacology & Toxicology, 97(5), 261–275.

DOI: 10.1002/asi

White, H.D., & McCain, K.W. (1997). Visualization of literatures. Annual Review of Information Science and Technology, 32, 99–168.
Whitley, K.M. (2002). Analysis of SciFinder Scholar and Web of Science citation searches. Journal of the American Society for Information Science and Technology, 53(14), 1210–1215.
Wleklinski, J.M. (2005). Studying Google Scholar: Wall to wall coverage? Online, 29(3), 22–26.
Zhao, D.Z., & Logan, E. (2002). Citation analysis using scientific publications on the web as data source: A case study in the XML research area. Scientometrics, 54(3), 449–472.

DOI: 10.1002/asi