Assessing the Scholarly Impact of Information Studies: A Tale of Two Citation Databases—Scopus and Web of Science

Lokman I. Meho
1320 East 10th Street, School of Library and Information Science, Indiana University, Bloomington, IN 47405. E-mail: meho@indiana.edu

Cassidy R. Sugimoto
School of Information and Library Science, University of North Carolina at Chapel Hill, 100 Manning Hall, CB #3360, Chapel Hill, NC 27514-3360. E-mail: csugimoto@email.unc.edu

This study uses citations, from 1996 to 2007, to the work of 80 randomly selected full-time, information studies (IS) faculty members from North America to examine differences between Scopus and Web of Science in assessing the scholarly impact of the field focusing on the most frequently citing journals, conference proceedings, research domains and institutions, as well as all citing countries. Results show that when assessment is limited to smaller citing entities (e.g., journals, conference proceedings, institutions), the two databases produce considerably different results, whereas when assessment is limited to larger citing entities (e.g., research domains, countries), the two databases produce very similar pictures of scholarly impact. In the former case, the use of Scopus (for journals and institutions) and both Scopus and Web of Science (for conference proceedings) is necessary to more accurately assess or visualize the scholarly impact of IS, whereas in the latter case, assessing or visualizing the scholarly impact of IS is independent of the database used.

Introduction

Citations, or references found in scholarly literature, are frequently used to: (a) assess or map the relative impact and perceived utility of research; (b) study the emergence and diffusion of new research ideas, journals, and communities; and (c) identify the extent of scholarly communication across research domains (Cronin & Meho, 2008; Garfield, 1998; Grover, Ayyagari, Gokhale, Lim, & Coffey, 2006; Lockett & McWilliams, 2005; White, 2004). These assessments or maps have implications for studying: (a) interdisciplinary research (Leydesdorff, 2007a, 2007b, 2007c; Liu & Wang, 2005; Tang, 2004); (b) the development of departmental structures (Borgman & Rice, 1992); (c) the evolution of journals (Goldstone & Leydesdorff, 2006); (d) the discovery of new knowledge (Chen, Zhang, Zhu, & Vogeley, 2007); (e) the history and growth of scientific fields (Li, Lin, Chen & Roco, 2007; Perry & Rice, 1998); (f) sociocognitive networks of scholars (Cronin, 2007); and (g) the impact of research carried out by different scientists, schools, and countries (Benamer & Bakoush, 2009; Hendrix, 2008; Jeang, 2008; Sánchez & Benn, 2004). This range of implications makes it imperative that appropriate citation databases and collection methods are utilized. The use of inappropriate citation databases or incomplete data risks developing inaccurate assessments that may generate distorted representations of the scholarly impact of research and lead to erroneous conclusions and decisions.

Researchers using citations to assess or map the scholarly communication and impact of authors, projects, programs, journals, countries, and so on, have often relied on data exclusively from three Institute for Scientific Information (ISI; currently Thomson Reuters) databases: Arts & Humanities Citation Index, Science Citation Index, and Social Sciences Citation Index. Reliance on these databases was partially due to the fact that these databases (which currently can be searched individually or together through Web of Science) were, for decades, the only comprehensive citation data sources that allowed large-scale analyses of, and comparisons among, authors, journals, disciplines, departments, universities, and countries (Meho, 2007; Moed, 2005). The emergence of Elsevier's Scopus database in late 2004, however, has raised many questions regarding (a) the validity of findings based exclusively on data from Web of Science, (b) the value and necessity of using multiple citation databases for assessing the scholarly impact of research, and

Received April 24, 2009; revised May 31, 2009; accepted June 1, 2009

 • Published online 24 August 2009 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/asi.21165

(c) the appropriateness of using Scopus as an alternative source of citation data to Web of Science. These questions arise because Scopus covers more than 15,000 titles in comparison to approximately 9,000 in Web of Science (for more details on both databases, see Gavel & Iselid, 2008; Meho & Rogers, 2008). Although numerous studies have compared the bibliographic and citation coverage of the two databases with each other (see below), none of these studies have focused on the implications of the use of one database rather than the other on assessing the scholarly impact of a research domain—a popular research area among scientometricians as evidenced by the increasing number of published literature on the topic (see Bar-Ilan, 2008).

Examples of studies that compared the bibliographic coverage of Scopus and Web of Science—but not the implications of using one rather than the other on assessing the scholarly impact of a research domain—include work by Norris and Oppenheim (2007), who used all but 720 of the journal articles submitted for the purpose of the United Kingdom 2001 Research Assessment Exercise in the social sciences ($n = 33{,}533$), as well as the list of 2,800 journals indexed in the *International Bibliography of the Social Sciences*, to assess the coverage of CSA Illumina, Google Scholar, Scopus, and Web of Science. They found that Scopus provides the best coverage of social science literature from among these data sources and concluded that Scopus could be used as an alternative to Web of Science as a tool to evaluate research impact in the social sciences. Gavel and Iselid (2008) compared the coverage of active journal titles in Scopus and Web of Science, finding that 84% of the Web of Science titles are indexed in Scopus, whereas only 54% of the Scopus titles are indexed in Web of Science. They identified 6,256 unique journals for Scopus and 1,467 unique journals for Web of Science. Their study, however, was admittedly based on journal lists provided by the database producers rather than manual searching of the databases to determine the breadth and depth of coverage of each title. Gorraiz and Schloegel (2008) compared the coverage of top pharmacology and pharmacy journals in Scopus and Web of Science. They found that all of the 100 highest impact Web of Science-covered journals in the field are indexed in Scopus and that Scopus covers some additional high impact journals not indexed by Web of Science. Similarly, López-Illescas, de Moya-Anegón, and Moed (2008) compared the coverage of oncology journals in Web of Science and Scopus. They found that all 126 Web of Science-covered oncology journals are indexed in Scopus, but that Scopus covers an additional 106 journals. These 106 additional oncology journals, however, tend to have much lower impact factors than those commonly covered by the two databases.

Examples of studies that compared the citation coverage of Scopus and Web of Science to determine whether using one database over another (or using both databases) significantly changes the ranking of scholars, publications, or topic retrieval include work by Bauer and Bakkalbasi (2005). They compared citation counts provided by Scopus, Google Scholar, and Web of Science for articles from the *Journal of the American Society for Information Science and Technology* published in 1985 and 2000. The results for 1985 articles were inconclusive; but, for 2000 articles, there were no significant differences between Scopus and Web of Science. Jacsó (2005) conducted several tests comparing Google Scholar, Scopus, and Web of Science, 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 there is considerable overlap between Scopus and Web of Science. He also found many unique documents in each source. Bar-Ilan, Levene, and Lin (2007) compared the rankings of the publications of 22 highly-cited Israeli researchers as measured by the citation counts in Google Scholar, Scopus, and Web of Science. Their results showed high similarity between Scopus and Web of Science and lower similarities between Google Scholar and the other databases.

Meho and Yang (2007) used citations to more than 1,400 works by 25 library and information science faculty to examine the effects of additionally using Scopus and Google Scholar on the citation counts and rankings of these faculty members as measured by Web of Science. They found that the addition of Scopus citations to those of Web of Science significantly altered the relative ranking of those faculty in the middle of the rankings, and that the use of Scopus and Google Scholar, in addition to Web of Science, reveals a more comprehensive and complete picture of the extent of the scholarly relationship between library and information science and other fields. Meho and Rogers (2008) examined differences between Scopus and Web of Science in the citation counting, citation ranking, and h-index of 22 top human-computer interaction (HCI) researchers from a large British Interdisciplinary Research Collaboration project. Their results showed that Scopus provides significantly more coverage of HCI literature than Web of Science, primarily due to broader coverage of conference proceedings, and that Scopus generates significantly higher h-index scores as well as different maps of citation networks of individual scholars from those generated by Web of Science. The authors concluded that Scopus can be used as a sole data source for citation-based research and evaluation in HCI. More recently, Levine-Clark and Gil (2009) examined the citation coverage in Web of Science, Scopus, and Google Scholar for a set of 15 business and economics journals. They concluded that all three data sources should be used to get a more complete picture of the scholarly impact of an article in these journals.

Most of the studies reviewed here indicate that the question of whether to use Scopus and/or Web of Science as part of a research assessment exercise may be domain-dependent and that more in-depth studies are needed to verify which database(s) is/are appropriate for what research domains and why. To investigate whether this is the case, this study uses a random sample of 80 full-time information studies (IS) faculty members from North America to examine differences between Scopus and Web of Science in assessing the scholarly impact of the field. It could be that there are marked differences in citation coverage between the two databases

DOI: 10.1002/asi

because of both the diversity of places where these faculty members publish and are cited (for more on this, see Meho & Yang, 2007) as well as their broad base—many with backgrounds in disciplines other than IS (for more on this, see Cronin & Meho, 2008).

Users of citations for examining intellectual trade among research domains, the development of departmental structures, the evolution of journals, discovery of new knowledge, the growth of scientific fields, the socio-cognitive networks of scholars, and the impact of research carried out by different entities (e.g., scholars, groups, departments, universities, countries) need to know: (1) whether literature coverage in one database (i.e., Scopus or Web of Science) generates significantly different results from that of the other database and, if it does, how that influences the assessment of certain scholars, disciplines, or topic domains; and (2) whether use of both databases is necessary to generate accurate assessments of scholarly impact.

Examining differences in research impact assessment between Scopus and Web of Science is important because it allows one to compare the consistency of the databases in such assessments. If differences are found between research domains, people who use citation analysis to demonstrate the relative impact and perceived utility of research, study the emergence and diffusion of new research ideas and communities, and identify the extent of scholarly communication across research domains will need to justify their choice of databases. Simply claiming that one database is more established or more comprehensive than the other would be insufficient.

Methods

To carry out the study, one of the authors developed a directory of all faculty members at American Library Association-accredited library and information science programs, based on information listed on the programs' Web sites as well as feedback from faculty members from these programs (Meho, 2007). The directory, which provides the names, status, affiliation, contact information, and Web site addresses (if available) of all IS full-time faculty members in North America, was then used to (a) generate a stratified random sample by rank for the study and (b) download publication information for the 80 sample members (30 assistant professors, 26 associate professors, and 24 professors). Where needed, publication information was requested from and provided by the sample members. While collecting citation data, it was discovered that 26% of the sample members did not provide or have online a complete list of their publications; this was determined when citations were occasionally found to relevant items not listed on the curriculum vita of the sample members. For these sample members, bibliographic searches were conducted in Inside Conferences, INSPEC, Library and Information Abstracts, Science Library Literature and Information Science, Social Sciences Citation Index, and WorldCat—databases considered by Meho and Spurgin (2003) most comprehensive in covering IS literature—to complete and verify publication information.

To collect citation data from Scopus, we used three searching methods: author search, search through the "More" tab, and exact match searching. In the first method, we identified for each sample member all of his or her publications in the database and retrieved all the citations to these publications. In the second method, we used the "More" searching/browsing feature to display, select, and collect citation data to items not found through or covered by the author search method (examples of these items are books, chapters in books, technical reports, dissertations, and journal articles and conference papers not indexed by the database). In the exact-match search method, we used the title of an item as a search statement (e.g., *Information Seeking in Electronic Environments*) and tried to locate an exact match in the cited "References" field of the indexed records. In cases where the title was too short or ambiguous to refer to the item in question, we used additional information as keywords (e.g., the first author's last name) 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 may increase the possibility of missing some relevant citations due to typing or indexing errors. The exact-match search method was most practical for authors with common last names, whereas the combination of author and "More" search methods was more practical for authors with less common last names. In Web of Science, we used the "Cited Reference Search" method to identify all relevant citations in the database. Where necessary, we used different permutations and search strategies to ensure that we captured all relevant citations.

The citation data, which were collected in March 2007 and updated and verified for accuracy in April 2009, were entered into an Excel spreadsheet and Access database. They were coded as follows: first author; source (e.g., journal, conference name); subject category; document type (e.g., journal article, review article, conference paper); reference type (e.g., journal vs. conference proceeding); publication year; language; institutional affiliation and country of the correspondence author; and the source used to identify the citation.[1] Virtually all citations were from refereed sources. Data from *ISI Proceedings*, which became searchable from within Web of Science in 2009, were not included in this study because it required a separate subscription.

Approximately 2% of the citations did not have country and institutional affiliation information. The Web was used to identify missing information. Because some journal and conference names are not entered in the same way in both Scopus

[1]To make sure that citing papers were not overlooked because of searching errors, and to account for database updates, in April 2009, we searched for all of the citing papers that were missed by each database. For example, if a citation was found in Scopus but not Web of Science, we conducted bibliographic searches in Web of Science to see if the item were in fact indexed in the database. When the bibliographic record of any of these missed citing papers was found in one of the two databases, we counted them as citing papers.

DOI: 10.1002/asi

TABLE 1. Distribution of citing papers by source type.

Source type	Scopus Count	Scopus Percent	WoS Count	WoS Percent	Total Count	Total Percent
Journals	5,288	74.4%	4,451	84.5%	5,501	73.4%
Conference proceedings	1,805	25.5%	807	15.3%	1,978	26.4%
Books in series	10	0.1%	11	0.2%	18	0.2%
Total	7,103	100.0%	5,269	100.0%	7,497	100.0%

and Web of Science (e.g., *Information Research* is indexed as *Information Research* in Scopus whereas it is *Information Research-An International Electronic Journal* in Web of Science), all such instances were manually standardized. In cases where a citing source had changed its name, the citations were merged under their most recent respective name (e.g., citations found in the *Journal of the American Society for Information Science* were listed under the journal's more current title, the *Journal of the American Society for Information Science and Technology*).

Results

Because Scopus provides citation data only since 1996 (in comparison to Web of Science's 1900 on), all the analyses reported here are limited to citing papers published between January 1996 and March 2007—the time data was collected for this study. Results show that the study sample has been cited in a total of 7,497 papers, including 5,501 (73.4%) from journals, 1,978 (26.4%) from conference proceedings, and 18 (0.2%) from items in series. Of these 7,497 papers, 7,103 (95%) were identified by Scopus and 5,269 (70%) by Web of Science (see Table 1 for a breakdown of citing papers by source type). We were not surprised by the higher number of citing papers found through Scopus considering that it covers many more titles than Web of Science (see Gavel & Iselid, 2008; Meho & Rogers, 2008). Results also show that 30% (2,228) of the citing papers were found uniquely in Scopus, whereas only 5% (394) of the citing papers were found uniquely in Web of Science. Sixty-five percent (4,875) of the citing papers were found in both databases.

The seemingly high number of citations found in only one of the databases (2,622 or 35% of the total number of citing papers) may suggest that both should be used for accurate assessment of the scholarly impact of IS research. Analyses that are more detailed were needed, however, to investigate whether these unique citing papers result in different assessments of scholarly impact of IS in terms of most frequently citing sources, research domains, institutions, and countries. If the unique citing papers in each database do not produce significant differences in assessing the scholarly impact of IS, using such categories as most frequently citing sources, research domains, institutions, and countries, then it is reasonable to conclude that assessing the scholarly impact of IS is largely independent of the citation database used. The remainder of this paper explores whether this is the case in the aforementioned categories.

Assessing Scholarly Impact of IS by Source Title: Journals and Conferences

Overall, the study sample has been cited in 1,786 sources (1,249 journals, 529 conferences, and 8 series). Results show that the citing papers identified through Scopus were found in 1,194 journals, 478 conferences, and 5 series, whereas those identified through Web of Science were found in 919 journals, 264 conferences, and 6 series. In both Scopus and Web of Science, more than 70% of the citing papers were found among the top 20% most citing sources with a very long tail of less frequently citing sources, reflecting a Bradford's law of scattering in citations—where a small number of sources account for the majority of citations and a large number of sources account for relatively few (Bradford, 1985).

Predicated on the assumption that the most frequently citing sources are among the most important or valuable to a research field or domain (Nisonger, 2007), we focused on examining potential assessment differences between Scopus and Web of Science in terms of the identity and ranking of the top 25 most frequently citing journals and top 10 most frequently citing conference proceedings. We also examined whether one or both databases should be used in such an assessment exercise. As can be seen in Table 2, results show that, with the exception of one title—*Journal of Educational Computing Research*—the top 25 most frequently citing journals are the same in both databases. There were, however, considerable differences in the citation counting and ranking of three journals—*Journal of Computer-Mediated Communication*, *Information Research*, and *Journal of Educational Computing Research*—,which resulted from differences in coverage in the two databases. All three journals are covered in Scopus from 1996, whereas in Web of Science, *Journal of Computer-Mediated Communication* is covered from 2005, *Information Research* from late 2002, and *Journal of Educational Computing Research* from 1987 to 2001.[2] As a result, in a research assessment exercise illustrating the scholarly impact of IS on scholarly journals within and outside the field, *Journal of Computer-Mediated Communication*, *Information Research*, and *Journal of Educational Computing Research* would be considerably more visible in Scopus, where they contribute 72, 62, and 37 citing papers, respectively; in a Web of Science-derived map, these sources contribute 46, 44, and 24 citing papers, respectively. Moreover, *Journal of Computer-Mediated Communication*, *Information Research*, and *Journal of Educational Computing Research* rank 7th, 12th, and 21st in Scopus, and in Web of Science, they rank 15th, 17th, and 36th, respectively. Had the analysis included journals beyond the 25 most frequently citing ones, more instances of considerable variation between Scopus and Web

[2]Table 2 shows additional minor differences in the number of citing papers for the same journal in both databases, with an advantage of Web of Science over Scopus. These differences are due to what Meho and Yang (2007) and Meho and Rogers (2008) referred to as "database errors" (e.g., incomplete coverage of some sources, missing cited references information from some bibliographic records, lack of coverage of certain types of documents [such as book reviews], and incomplete coverage of editorial material).

DOI: 10.1002/asi

TABLE 2. Top 25 most frequently citing journals.

	Number of citing papers			Citation ranking		
	Scopus	WoS	Both	Scopus	WoS	Both
Journal of the American Society for Information Science and Technology	332	338	342	1	1	1
Information Processing & Management	175	176	177	2	2	2
International Journal of Human Computer Studies	*106*	*106*	*106*	*3*	*3*	*3*
Annual Review of Information Science and Technology	81	82	82	4	4	4
Library & Information Science Research	80	80	81	5T	5T	5
Interacting With Computers	*80*	*80*	*80*	*5T*	*5T*	*6*
Journal of Computer-Mediated Communication	*72*	*46*	*72*	*7*	*15*	*7*
Journal of Documentation	65	69	70	9	8	8T
Library Trends	66	70	70	8	7	8T
Information Research	62	44	66	12	17	10T
Scientometrics	63	66	66	11	9	10T
Journal of Information Science	64	64	64	10	10	12
Computers in Human Behavior	*60*	*60*	*61*	*13*	*11*	*13*
Behaviour & Information Technology	*50*	*51*	*52*	*14*	*12*	*14*
Library Quarterly	49	50	51	15	13	15
Journal of the Medical Library Association	48	49	49	16	14	16
Computers & Education	*42*	*45*	*45*	*17*	*16*	*17*
Online Information Review	41	43	43	18	18	18
Electronic Library	40	40	41	19	19	19
Journal of Visual Languages and Computing	*38*	*38*	*38*	*20*	*20*	*20*
Journal of Educational Computing Research	*37*	*24*	*37*	*21*	*NR*[a]	*21*
International Journal of Human-Computer Interaction	*36*	*35*	*36*	*22*	*21T*	*22T*
Journal of Academic Librarianship	35	35	36	23	21T	22T
Communications of the ACM	*34*	*34*	*35*	*24*	*23*	*24*
Aslib Proceedings	33	33	34	25	24	25

[a] *Journal of Educational Computing Research* ranks 36th most frequently citing journal in Web of Science. The 10 journals in italics are from outside the field, including communication, computer science, educational technology, human-computer interaction, and information systems. (See discussion below on assessing scholarly impact of IS by research domain.) Note that by the end of 2006, Scopus filled in almost all of the gaps in its coverage of post-1995 issues of the 15,000 titles it indexes (see Goodman & Deis, 2007).

of Science would have appeared (with more complete or more accurate results from Scopus). Accordingly, it was concluded that to assess the scholarly impact of North American IS faculty research on IS and non-IS journals more accurately, Scopus should be used, especially if analysis is limited to citations from 1996 on.

Unlike the case of journals, when examining the top 10 most frequently citing conference proceedings overall and in each database, results show that Scopus and Web of Science share only one title—the *Proceedings of the Annual Meeting of the American Society for Information Science and Technology* (see Table 3). These results suggest that each database would produce an almost entirely different assessment of the scholarly impact of IS research based on the most frequently citing conference proceedings. The assessment based on Scopus data would show the prominence of ACM and IEEE conference proceedings (with a focus on digital libraries, human-computer interaction, and information retrieval), whereas the one based on Web of Science would show the prominence of conference proceedings produced by different professional organizations and published in the series *Lecture Notes in Computer Science* (with a focus on computer science and engineering). Even though Scopus covers 95% of the papers citing the study sample as well as the top 10 most frequently citing conference proceedings, Web of Science is still necessary to generate a more complete list of conferences influenced by North American IS faculty research. This is primarily because Web of Science provides more complete coverage of such frequently citing conference proceedings as *Research and Advanced Technology for Digital Libraries*, *Advances in Information Retrieval*, and *Advances in Case-Based Reasoning* (see Table 3).

Assessing Scholarly Impact of IS by Research Domain

Using Web of Science to examine intellectual trade in IS, Cronin and Meho (2008) found that (during 1997 and 2006) approximately 65% of all the papers that cited IS journals and conference proceedings came from non-IS sources. In comparison, this study found that of the 7,497 papers that have cited the study sample, 73% (5,479) came from outside the field (this difference is largely due to differences in literature coverage between Scopus and Web of Science). These non-IS citing papers were found in 1,703 journals and conference proceedings; the remaining 2,018 citing papers came from 83 IS journals and conference proceedings.

To determine whether one database would generate a significantly different picture or illustration than the other, in terms of scholarly impact of IS on other fields, it was necessary to classify the 1,703 non-IS citing journals and conference proceedings by subject category. Because it was effectively impractical to assign subject categories to all of

DOI: 10.1002/asi

TABLE 3. Top 10 most frequently citing conference proceedings in each database.

	Number of citing papers			Citation ranking		
	Scopus	WoS*	Both	Scopus	WoS	Both
Proceedings of the Annual Meeting of the American Society for Information Science and Technology (ASIST)	91	71	93	1	1	1
CHI: ACM Conference on Human Factors in Computing Systems	75	not covered	75	2	NA	2
Annual Hawaii International Conference on System Sciences (IEEE)	55	not covered	55	3	NA	3
Proceedings of SPIE - The International Society for Optical Engineering	49	not covered	49	4	NA	4
ACM SIGIR Conference on Research and Development in Information Retrieval	47	not covered	47	5	NA	5
ACM IEEE International Conference on Digital Libraries, JCDL	40	not covered	40	6	NA	6
ACM International Conference on Information and Knowledge Management	37	not covered	37	7	NA	7
Research and Advanced Technology for Digital Libraries (LNCS)	14	36	36	14T	2	8
ACM SIGKDD International Conference on Knowledge Discovery and Data Mining	28	not covered	28	8	NA	9
Advances in Information Retrieval (LNCS)	11	27	27	22T	3	10
IEEE International Conference on Systems, Man and Cybernetics	26	not covered	26	9	NA	11
ACM Symposium on Applied Computing	25	not covered	25	10	NA	12
Advances in Case-Based Reasoning (LNCS)	2	19	19	140T	4	14
Knowledge-Based Intelligent Information and Engineering Systems (LNCS)	6	13	13	46T	5	19T
Advances in Artificial Intelligence (LNCS)	6	12	12	46T	6T	22T
Computational Linguistics and Intelligent Text Processing (LNCS)	5	12	12	56T	6T	22T
Intelligent Tutoring Systems (LNCS)	2	12	12	140T	6T	22T
Image and Video Retrieval (LNCS)	4	11	11	76T	9T	29T
Intelligent Data Engineering and Automated Learning (LNCS)	2	11	11	140T	9T	29T

Note. WoS indicates Web of Science. LNCS refers to the series *Lecture Notes in Computer Science*, which is published by Springer. Web of Science ceased covering *LNCS* in 2006. *LNCS* and many of the conference proceedings indexed in Scopus are covered in *ISI Proceedings*, which was not used in this study.

these 1,703 non-IS sources, analysis was limited to the top 200 most frequently citing sources, which accounted for 2,893 or 53% of the 5,479 non-IS citing papers. As did Cronin and Meho (2008), we used citation analysis (relying on both Scopus and Web of Science) and our domain knowledge to assign each source to a subject domain. We did not exclusively use Web of Science's journal subject classification because its categories were too broad. To illustrate our method, an analysis of the titles that have cited *MIS Quarterly*, for example, revealed that it has been cited mostly by information systems journals, followed by computer science. In this case, we classified *MIS Quarterly* as an information systems journal. We applied the same approach to all of the 200 most frequent non-IS citing sources. This subject classification resulted in a total of 33 citing research domains, 10 of which each had approximately 1% or more of the total number of citing papers (see Figure 1).

As can be seen in Figure 1, results show only marginal differences between Scopus and Web of Science in assessing the scholarly impact of North American IS faculty research on the top 10 most frequently citing research domains. In fact, the assessments produced by both databases are almost identical with human-computer interaction, computer engineering, information systems, and educational technology, in this order, being the main importers of the IS literature in question, exactly the same research domains as those found by Cronin and Meho (2008) for the period 1997–2006. Furthermore, the results of this study show that the merger of data from both databases would not significantly alter the assessment of scholarly impact of North American IS faculty research on other domains, suggesting that such assessment exercises for IS faculty research in North America is largely independent of the citation database used and that the use of one database is sufficient in producing an accurate picture of scholarly impact of IS by research domain.

Assessing Scholarly Impact of IS by Institution

Assessing the scholarly impact of a research domain or a group of scholars by the institutional affiliation of citing papers or authors is a rare but an important exercise. It gives an indication of both (a) the relative scholarly influence or impact the research domain or the group of scholars has (or has had) on these institutions and (b) the relative scholarly productivity of the citing institution in research areas relevant to the subject domain or group of scholars in question—recall that 73% of all the citing papers to the work of the study sample come from non-IS sources. Results here show that the study sample has been cited by authors from 1,764 institutions, including universities, government agencies, corporate organizations, and libraries, among others. Of these institutions, 20% accounted for 5,232 (70%) of the total number of citing papers. It was not surprising to find out that Scopus accounted for more institutions (1,711) than Web of Science (1,378), simply because of its coverage of more sources.

As in the case of citing sources (i.e., journals and conference proceedings), when assessing the scholarly impact of a research domain by the institutional affiliation of citing papers or authors, one is more likely to focus on the top 25 or so institutions with the largest number of citing

DOI: 10.1002/asi

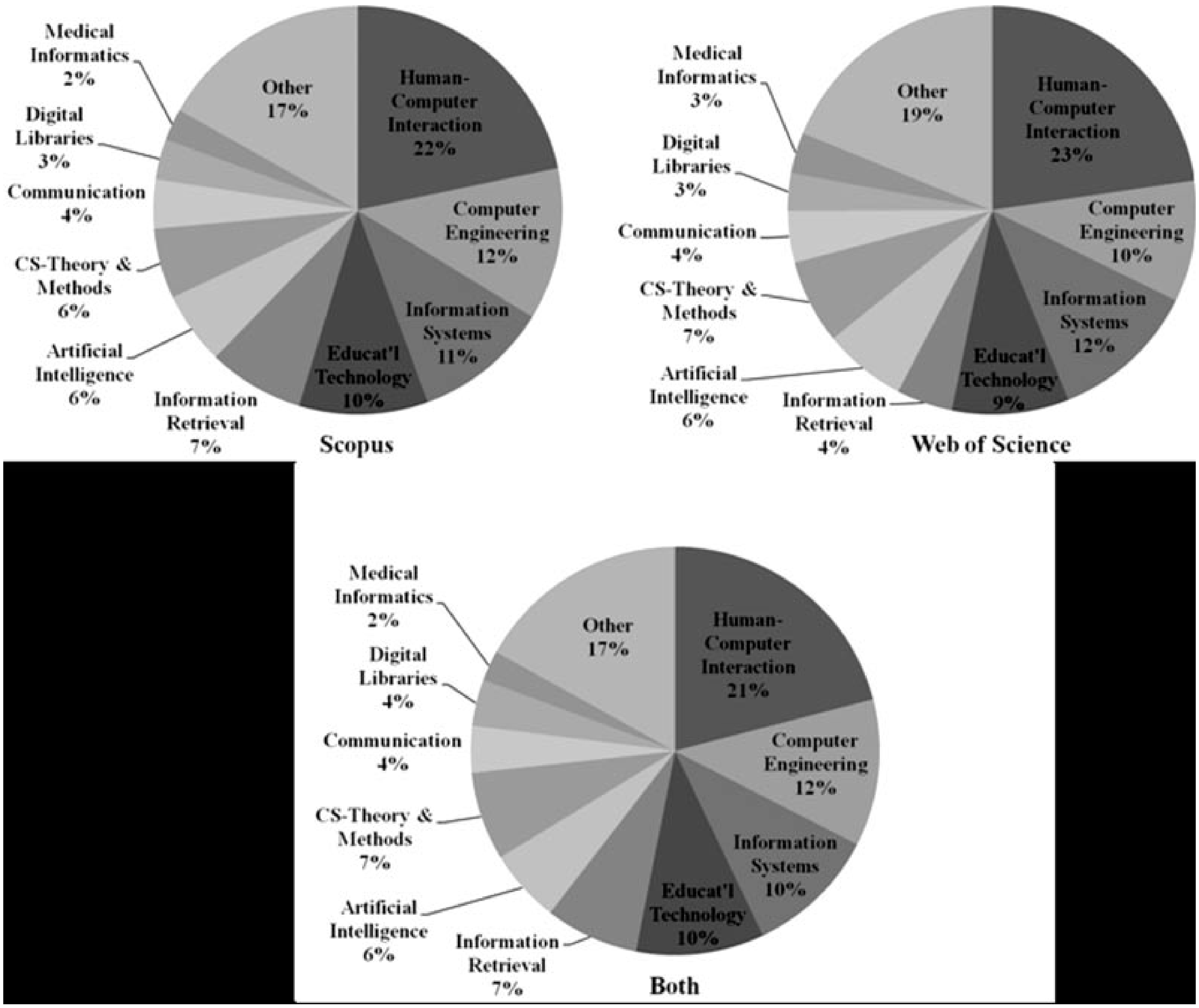

FIG. 1. Distribution of citing papers by research domain of top 10 importers.

papers or authors (citing institutions henceforth). After all, the most frequently citing institutions can be considered the main beneficiaries of, or the most influenced by, the scholarly products of the subject domain or group of scholars in question. Examining the scholarly impact of the study sample on the 25 most frequently citing institutions shows significant differences between Scopus and Web of Science for institutions ranking 7th through 25th—with Spearman rank order correlation coefficient of .523 at the 0.01 level for a two-tailed test (see Table 4). An assessment based on Web of Science illustrating the scholarly impact of the study sample by the top 25 most frequently citing institutions would omit three institutions (IBM, Microsoft, and Virginia Polytechnic Institute and State University) that could have been observed had the assessment been based on citation data from Scopus. These two assessments also show considerable ranking differences of most frequently citing institutions by a minimum margin of six points in eight cases out of the 19 institutions ranking 7th through 25th (11 cases with a ranking difference of five points or more). Perhaps most importantly, a map based exclusively on Web of Science citation data would miss the influence of North American IS faculty research on industry. Finally, adding citation data from Web of Science to those of Scopus would not make any significant differences to the assessment of the top 25 most citing institutions. Accordingly, as in the case of most frequently citing journals, these results suggest that when it comes to assessing scholarly impact of a research domain by the most frequently citing institutions, care should be taken in the selection of the database because it considerably affects the distribution and rankings of these institutions.

Assessing Scholarly Impact of IS by Country

Citations are often used to examine, illustrate, and make comparisons of the international impact of research carried out by journals, scholars, groups, departments, institutions, and countries (Van Raan & Van Leeuwen, 2002; Zainab, 2008). Results here show that the study sample has been cited by authors from 71 countries, which were divided into

DOI: 10.1002/asi

TABLE 4. Top 25 most frequently citing institutions.

Institution	No. of citing papers			Rank		
	Scopus	WoS	Both	Scopus	WoS	Both
University of North Carolina at Chapel Hill	100	77	101	1	2	1
Indiana University Bloomington	95	80	99	2	1	2
University of Maryland, College Park	70	61	77	5	3	3T
University of Toronto	77	54	77	3	5	3T
University of Illinois at Urbana-Champaign	72	55	75	4	4	5
University of California, Los Angeles	61	52	65	6T	6	6
Carnegie Mellon University	57	35	62	8T	18	7T
University of Arizona	61	50	62	6T	7	7T
University of Western Ontario	55	46	58	10	8	9
University of Washington	57	39	57	8T	14	10
Nanyang Technological University	51	42	56	13	10T	11T
Rutgers State University	53	44	56	11T	9	11T
IBM	53	24	53	11T	(38T)	13
Pennsylvania State University, University Park	49	37	52	14	15T	14
University of Michigan	45	40	50	17T	12T	15
University of Minnesota	48	33	49	15T	22	16T
University of Wisconsin, Milwaukee	48	42	49	15T	10T	16T
Drexel University	45	31	46	17T	25T	18T
Open University	44	34	46	19T	19T	18T
University of Pittsburgh	44	37	46	19T	15T	18T
Microsoft	44	19	45	19T	(46T)	21
McGill University	44	40	44	19T	12T	22
Syracuse University	43	36	43	23	17	23
Stanford University	38	27	42	(26T)	(33T)	24
Virginia Tech	39	29	41	24T	(29)	25

Note. WoS indicates Web of Science.

11 geographical regions (based on United Nations classification), excluding Canada and the United States—the two countries in which the study sample members work (see Figure 2). Slightly more than half (50.5%) of the 7,497 citing papers came from Canada and the United States; the remaining 49.5% came from authors working in other parts of the world, primarily the United Kingdom and Ireland (12.9%), Western Europe (9.1%), Eastern Asia (8.0%), Southern Europe (5.3%), and Northern Europe (4.7%).

As Figure 2 shows, Scopus and Web of Science produce almost identical maps of the distribution of citations to the work of the study sample, and, as in the case of classification of citations by research domain, merging citation data from both databases do not make any significant difference. These results suggest that for larger aggregates of citation data (e.g., citation distribution by research domain and country), assessment of scholarly impact of research is largely independent of the citation database used. In contrast, for smaller aggregates of citation data (e.g., citation distribution by source and institution), assessment of scholarly impact of research can be influenced, sometimes significantly as was the case with conference proceedings in this study, by the citation database used.

Conclusion

This study uses citations to the work of 80 randomly selected full-time, IS faculty members from North America to examine differences between Scopus and Web of Science in assessing the scholarly impact of the field and focusing on the most frequently citing journals, conference proceedings, research domains, and institutions, as well as all citing countries. Results showed that when analysis is based on smaller citing entities, such as journals, conference proceedings, and institutions within and beyond the field, the two databases produce considerably different results, whereas when analysis is based on larger citing entities, such as research domains and countries, the two databases produce very similar pictures of scholarly impact. In the former case, the use of Scopus (for journals and institutions) and both Scopus and Web of Science (for conference proceedings) is necessary to more accurately assess or visualize the scholarly impact of IS, whereas in the latter case, assessing or visualizing the scholarly impact of IS is independent of the database used

The findings reported here raise questions regarding the validity and accuracy of studies that are or were carried out without first verifying whether the use of Scopus, Web of Science, or both is necessary in assessing and/or mapping research impact. The need to use one or both citation databases in such exercises will likely vary among research domains, necessitating further investigation. Such work is important because the use of inappropriate citation databases or incomplete data risks both developing inaccurate assessments and generating distorted representations of the scholarly impact of research, which, in turn, can lead to erroneous conclusions and decisions. Users of citations—for assessing

DOI: 10.1002/asi

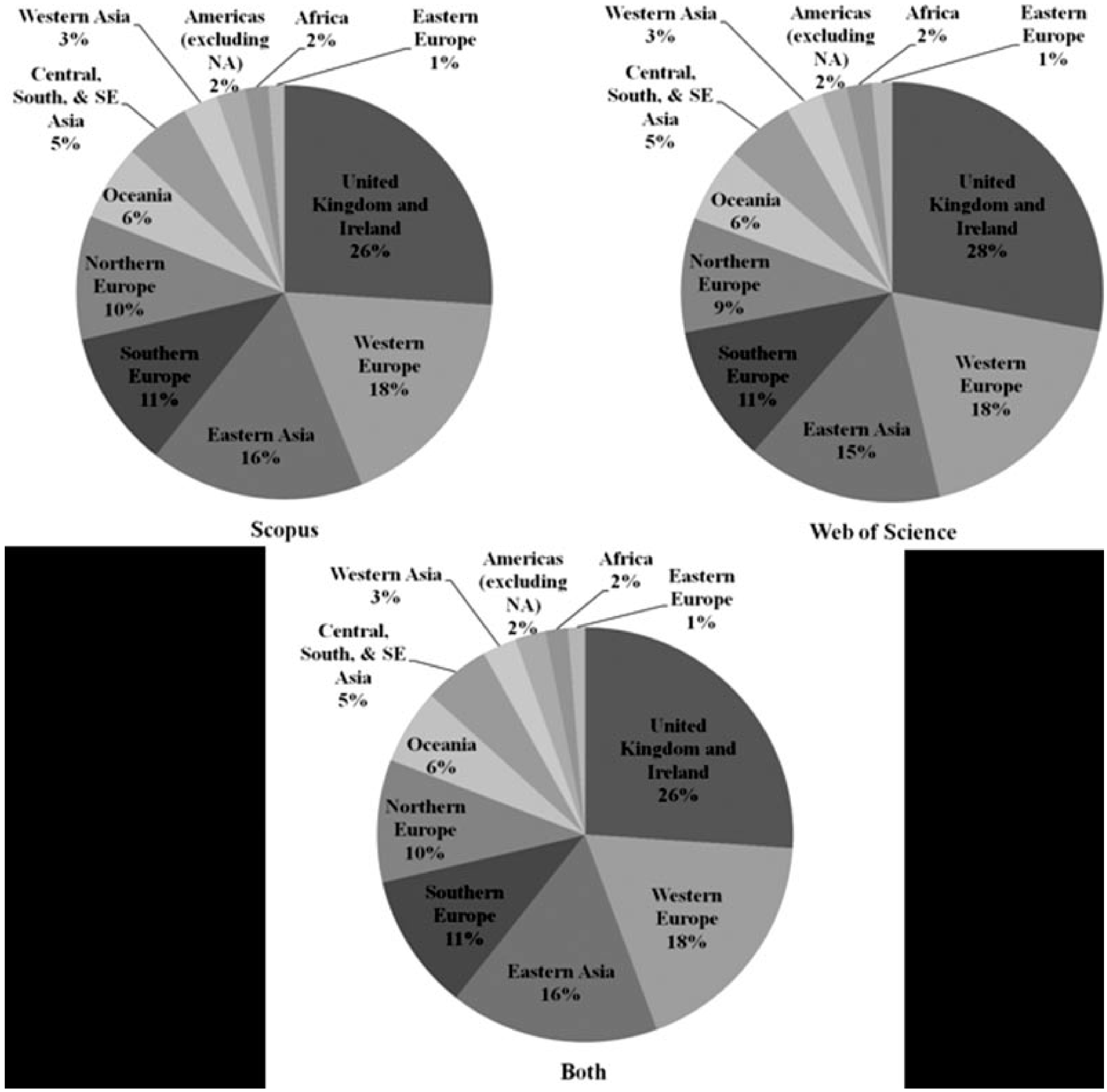

FIG. 2. Distribution of citing papers by region (excluding Canada and the United States).

or visualizing intellectual trade among research domains and examining the development of departmental structures, evolution of journals, discovery of new knowledge, growth of scientific fields, socio-cognitive networks of scholars, and the impact of research carried out by different authors, groups, departments, institutions, and countries—must make it a goal to first determine the appropriate citation database or databases to use in such exercises and explain why they have made their choice before carrying out their research and disseminating their findings and conclusions.

Acknowledgment

We are very grateful to Debora Shaw and two anonymous referees for their valuable comments on the paper.

References

Bar-Ilan, J. (2008). Informetrics at the beginning of the 21st century—A review. Journal of Informetrics, 2(1), 1–52.

Bar-Ilan, J., Levene, M., & Lin, A. (2007). Some measures for comparing citation databases. Journal of Informetrics, 1(1), 26–34.

Bauer, K., & Bakkalbasi, N. (2005). An examination of citation counts in a new scholarly communication environment. D-Lib Magazine, 11(9). Retrieved April 10, 2009, from http://www.dlib.org/dlib/september05/bauer/09bauer.html.

Benamer, H.T.S., & Bakoush, O. (2009). Arab nations lagging behind other Middle Eastern countries in biomedical research: A comparative study. BMC Medical Research Methodology, 9(26). Retrieved June 1, 2009, from http://www.biomedcentral.com/1471-2288/9/26

Borgman, C.L., & Rice, R.E. (1992). The convergence of information science and communication: A bibliometric analysis. Journal of the American Society for Information Science, 43(6), 397–411.

DOI: 10.1002/asi

Bradford, S.C. (1985). Sources of information on specific subjects. Journal of Information Science, 10(4), 173–180.

Chen, C., Zhang, J., Zhu, W., & Vogeley, M. (2007). Delineating the citation impact of scientific discoveries. In R.R. Larson (Ed.), Joint conference on digital libraries (pp. 19–28). Washington, DC.: ACM.

Cronin, B., & Meho, L.I. (2008). The shifting balance of intellectual trade in information studies. Journal of the American Society for Information Science and Technology, 59(4), 551–564.

Cronin, B., & Shaw, D. (2007). Peers and spheres of influence: Situating Rob Kling. The Information Society, 23(4), 221–233.

Garfield, E. (1998). From citation indexes to informetrics: Is the tail now wagging the dog? Libri, 48(1), 67–80.

Gavel, Y., & Iselid, L. (2008). Web of Science and Scopus: A journal title overlap study. Online Information Review, 32(1), 8–21.

Goldstone, R.L., & Leydesdorff, L. (2006). The import and export of Cognitive Science. Cognitive Science, 30(6), 983–993.

Goodman, D., & Deis, L. (2007). Update on Scopus and Web of Science. The Charleston Advisor, 8(3), 15–18.

Gorraiz, J., & Schloegl, C. (2008). A bibliometric analysis of pharmacology and pharmacy journals: Scopus versus Web of Science. Journal of Information Science, 34(5), 715–725.

Grover, V., Ayyagari, R., Gokhale, R., Lim, J., & Coffey, J. (2006). A citation analysis of the evolution and state of information systems within a constellation of reference disciplines. Journal of the Association for Information Systems, 7(5), 270–325.

Hendrix, D. (2008). An analysis of bibliometric indicators, National Institutes of Health funding, and faculty size at Association of American Medical Colleges medical schools, 1997–2007. Journal of the Medical Library Association, 96(4), 324–334.

Jacsó, P. (2005). 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 April 10, 2009, from http://www.ias.ac.in/currsci/nov102005/1537.pdf

Jeang, K.T. (2008). H-index, mentoring-index, highly-cited and highly-accessed: How to evaluate scientists? Retrovirology, 5(106). Retrieved June 1, 2009, from http://www.retrovirology.com/content/5/1/106

Levine-Clark, M., & Gil, E.L. (2009). A comparative citation analysis of Web of Science, Scopus, and Google Scholar. Journal of Business and Finance Librarianship, 14(1), 32–46.

Leydesdorff, L. (2007a). Environment and planning B: Planning and design as a journal: The interdisciplinarity of its environment and the citation impact. Environment and Planning B-Planning & Design, 34(5), 826–838.

Leydesdorff, L. (2007b). Mapping interdisciplinarity at the interfaces between the science citation index and the social science citation index. Scientometrics, 71(3), 391–405.

Leydesdorff, L. (2007c). Visualization of the citation impact environments of scientific journals: An online mapping exercise. Journal of the American Society for Information Science and Technology, 58(1), 25–38.

Li, X., Lin, Y.L., Chen, H.C., & Roco, M.C. (2007). Worldwide nanotechnology development: A comparative study of USPTO, EPO, and JPO patents (1976–2004). Journal of Nanoparticle Research, 9(6), 977–1002.

Liu, Z., & Wang, C.Z. (2005). Mapping interdisciplinarity in demography: A journal network analysis. Journal of Information Science, 31(4), 308–316

Lockett, A., & McWilliams, A. (2005). The balance of trade between disciplines. Do we effectively manage knowledge? Journal of Management Inquiry, 14(2), 139–150

López-Illescas, C., de Moya-Anegón, F., & Moed, H.F. (2008). Coverage and citation impact of oncological journals in the Web of Science and Scopus. Journal of Informetrics, 2(4), 304–316.

Meho, L.I. (2007). The rise and rise of citation analysis. Physics World, 20(1), 32–36.

Meho, L.I., & Rogers, Y. (2008). Citation counting, citation ranking, and h-index of human-computer interaction researchers: A comparison of Scopus and Web of Science. Journal of the American Society for Information Science & Technology, 59(11), 1711–1726.

Meho, L.I., & Spurgin, K.M. (2005). Ranking the research productivity of library and information science faculty and schools: An evaluation of data sources and research methods. Journal of the American Society for Information Science and Technology, 56(12), 1314–1331.

Meho, L.I., & Yang, K. (2007). Impact of data sources on citation counts and rankings of IS faculty: Web of Science versus Scopus and Google Scholar. Journal of the American Society for Information Science & Technology, 58(13), 2105–2125.

Moed, H.F. (2005). Citation analysis in research evaluation. Dordrecht, The Netherlands: Springer.

Nisonger, T.E. (2007). Journals in the core collection: Definition, identification, and applications. The Serials Librarian, 51(3–4), 51–73

Norris, M., & Oppenheim, C. (2007). Comparing alternatives to the Web of Science for coverage of the social sciences' literature. Journal of Informetrics, 1(2), 161–169.

Perry, C.A., & Rice, R.E. (1998). Scholarly communication in developmental dyslexia: Influence of network structure on change in a hybrid problem area. Journal of the American Society for Information Science, 49(2), 151–168.

Sánchez, S.F., & Benn, C.R. (2004). Impact of astronomical research from different countries. Astronomische Nachrichten, 325(5), 445–450.

Tang, R. (2004). Evolution of the interdisciplinary characteristics of information and library science. In L.S. & C.L. Barry (Eds.), Proceedings of the 67th ASIS&T Annual Meeting (pp. 54–63). Silver Spring, MD: American Society for Information Science and Technology.

Van Raan, A.F.J., & Van Leeuwen, T.N. (2002). Assessment of the scientific basis of interdisciplinary, applied research application of bibliometric methods in nutrition and food research. Research Policy, 31(4), 611–632.

White, H.D. (2004). Citation analysis and discourse analysis revisited. Applied Linguistics, 25(1), 89–116.

Zainab, A.N. (2008). Internationalization of Malaysian mathematical and computer science journals. Malaysian Journal of Library & Information Science, 13(1), 17–33.