The current issue and full text archive of this journal is available at
www.emeraldinsight.com/0001-253X.htm

Knowledge input for the domain of information science

A bibliometric and citation analysis study

Ming-yueh Tsay
Graduate Institute of Library, Information and Archival Studies, National Chengchi University, Taipei, Taiwan, Republic of China

Abstract

Purpose – The aim of this paper is to explore the knowledge input and the subject relationship with other disciplines for the domain of information science through a citation analysis, from the references of each article from 1998 to 2008, of four leading information science journals, *Journal of the American Society for Information Science and Technology*, *Information Processing and Management*, *Journal of Information Science*, and *Journal of Documentation*.

Design/methodology/approach – The Ulrich's Periodical Directory, Library of Congress Subject Heading, retrieved from the WorldCat and LISA database were used to identify the main class, subclass and subject of cited journals and books. The highly cited journals and books, the main classes and subclasses of cited journals and books in papers of the four journals, the highly cited subjects in journals and books of library and information science were identified and analyzed.

Findings – The study reveals that information science possessing strong self knowledge flow as these four source journals are also the four most cited. The class library science, book industries and trade, general information resources consists of a mainstream of knowledge flow into information science. The highly cited subjects of LIS journals encompass searching, online information retrieval, information work, subject indexing, World Wide Web, technical services, citation analysis, information seeking behavior, etc. The three most cited LCSH subjects on WorldCat of books are information storage and retrieval, information science, human-computer interaction, etc.

Originality/value – The knowledge inputs for information science include mainly information science itself and social sciences and general science as well. Moreover, there are minor inputs from various subjects.

Keywords Bibliometric study, Cited books, Cited journals, Subject analysis, Knowledge input, Information science journal, Journals, Information science, Information studies

Paper type Research paper

1. Introduction

As is well-accepted, information science is an interdisciplinary science evolving from the interaction of many other disciplines, such as mathematics, logic, linguistics, psychology, computer technology, operations research, the graphic arts, communications, library science, management, and other similar fields. Indeed, Borko (1968, p. 3) defined that information science is "a discipline that investigates the properties and behavior of information, the forces governing the flow of information, and the means of processing

This work was supported by grant NSC-97-2410-H-004-153-MY3 from the National Science Council, Taiwan, Republic of China. Bibliometric data collected and processed by Ju-Yu Shu, Graduate Institute of Library, Information and Archival Studies, National Chengchi University, Taiwan, and Bih-Ling Fang, Department of Information and Library Science, Tamkang University, Taiwan are very much appreciated.

Aslib Proceedings: New Information Perspectives
Vol. 65 No. 2, 2013
pp. 203-220

0001-253X
DOI 10.1108/00012531311314005

information for optimum accessibility and usability". Saracevic (1999, p. 1052) also revealed that "information science is interdisciplinary in nature", "is connected to information technology" and is "an active participant in the evolution of the information society with a strong social and human dimension, above and beyond technology". Through the study of the origin of information from various perspectives Saracevic also found that information science is related to other fields from several aspects, including historical, sociological, philosophical, technological, educational, and interdisciplinary. It is, therefore, of significant interest to explore the knowledge input to the domain of information science through the study of interaction among various disciplines.

Bibliometric techniques using citation analysis may be applied to study scholarly communication flow. For example, citations can be clustered to identify the flow of topics within and among disciplines. It also can be used to map relationships among documents, among journals or other channels of scholarly communications (Borgman, 1999, p. 118). Indeed, citation analysis is an important area of library and information science. From the studies of citation analysis, one can learn which scholars from which disciplines cite which articles? Which journals are cited more often? Which disciplines cite the journals of other disciplines? The results of citation analysis can serve many purposes, for example, to determine the impact of specific articles or journals on subsequent research and to document the interdisciplinary applicability of various journals (Desai, 2003; Harter, 1996).

To serve the purpose of this study investigating the knowledge flows into the information science, the present work conducts a bibliometric and citation analysis on four leading journals in this subject, namely, *Journal of the American Society for Information Science and Technology* (*JASIST*), *Information Processing and Management* (*IPM*), *Journal of Documentation* (*JOD*), and *Journal of Information Science* (*JIS*). These four journals have been recognized as general-purpose journals, which publish articles about and from most areas of the discipline related to information science as revealed in the scope statements of the journals. Based on the bibliometric and citation analysis on these four leading journals in the field of information science, the present study may help understanding the interactions among the disciplines relating to information science and further discovering the major kinds of knowledge flowing into the domain of information science.

1.1 Literature review

There have been some bibliometric studies on the cited reference of a particular journal in information science in the literature. Employing a variety of bibliometric methods, including publication and citation analyses, Bonnevie (2003) investigated a multifaceted portrait of the *Journal of Information Science,* focused on the last quarter of the twentieth century. The journal co-citation analysis shows that *JIS* is mainly co-cited with journals in the field of LIS. *Journal of Documentation, JASIS* and *Scientometrics* are the top three journals closest to *JIS*. Based on analyses of references in journal articles and journal co-citation analyses, Nebelong-Bonnevie and Frandsen (2006) proposed the journal citation identity (i.e. references per different referenced work) and journal citation image as two indicators for journal evaluation. They analyzed *Journal of Documentation* (*JOD*) by using the data of *Journal of Information Science* (*JIS*) and *Journal of the American Society for Information Science and Technology* (*JASIST*) as standard of reference and comparison. The results of the journal co-citation analysis indicated that *JASIST* and *JIS* were the two journals closest

to *JOD* and the image of *JOD* was influenced, especially, by *JASIST* and *IPM* with an upward tendency and to a less degree by *JIS*.

By analyzing the monographic references cited in the 1987-1990 issues of three information science journals (*IPM*, *JASIS*, *JOD*), DeHart (1992) identified the five most frequently cited authors, involving 20 different books, namely G. Salton, CM. van Rijsbergen, R. Schank, M. Kochen, and F. Machlup. From OCLC records, DeHart also identified the five subjects most appearing often as information storage and retrieval systems; artificial intelligence; discourse analysis; database management; and human-computer interaction. For complimenting subject distribution, DeHart employed Dewey main classes to analyze disciplinary distribution of 172 cited monographs and found the following results: 000: generalities (60 times cited); 500: natural science and mathematics (35); 300: social sciences (26); and 100: philosophy and psychology (17). The results demonstrated shifts in disciplinary emphasis in the information science field.

Based on 490 cited authors cited by at least two articles that published in the *Journal of the American Society for Information Science*, Persson (1994) conducted an author co-citation analysis to find the intellectual base of the field of information science. His study revealed that bibliometrics and information retrieval were two major branches of the tradition of information science. Within the bibliometrics subfield, citation analysis and bibliometric distributions were the two main groups. The information retrieval cluster was further subdivided into "hard" part dealing with algorithms and "soft" part on the user-system relation.

White and McCain (1998) explored the domain knowledge of the information science field for the years 1972-1995 by author co-citation analysis. Citation and co-citation data were collected from the databases of Social SciSearch via DIALOG as the objects of analysis and factor analysis and multi-dimensional scale (MDS) were used to visualize the specialty structure of and development on the information science discipline. The results of their factor analysis yield 12 factors as following in order of Pearson correlation coefficients (loadings) greater than 0.3: experimental retrieval, citation analysis, practical retrieval, bibliometrics, general library systems theory, user theory, scientific communication, OPAC, imported ideas, indexing theory, citation theory, communication theory.

Following the time period studied, Zhao and Strotmann (2008), using an author co-citation, examined the information science field defined in the same way as White and McCain's study for the years 1996-2005. The research results from two factor analyses (with an orthogonal rotation and with an oblique rotation) identified 12 factors that presenting the web impact on information science. These factors include user studies, citation analysis, experimental retrieval, webometrics, visualization of knowledge domains, science communication, users' judgments of relevance (situational relevance), information seeking and context, children's information searching behavior, metadata and digital resources, bibliometric models and distributions, structured abstracts (academic writing).

Astrom (2007) analyzed articles from 21 LIS journals covering the years 1990-2004 and references found in these articles as indexed in the Web of Science to determine what research topics have dominated LIS during the study period. He used co-citation analyses to study research fronts (defined as current and influential co-cited articles) and employed MDS and cluster analysis techniques to visualize relationship between LIS subjects. The research base was based on the 13,605 journal articles and their 221,586 references to 150,145 unique documents. The 66 most-cited documents that received 50 citations or more were selected for further analysis. His analysis was done

on a document level as opposed to an analysis on the author level. The co-citation map of LIS research base demonstrated two main areas: information-seeking and retrieval and informetrics. His cluster analysis resulted in eight clusters, i.e. experimental information retrieval (IR), IR/information search, IR/relevance, information seeking and use/cognitive IR, information seeking and use/information behavior, bibliometric mapping, bibliometric distributions, world wide web/webometrics.

1.2 Objectives

The literature review above reveals that most previous studies were on the bibliometric analysis on single journal or two or three journals of *JIS*, *JOD*, *IPM*, and *JASIS(T)*. According to Paisley (1990) and McCarthy (2000), *JASIST*, *IPM*, *JIS* and *JOD* broadly represent the information science field. Some studies deal with building domain knowledge structure by co-citation analysis based on authors or journals. However, subject analysis on the references cited had been seldom studied. Therefore, the main objective of the present study is to analyze the subjects of cited references in *JASIST*, *IPM*, *JOD* and *JIS* from 1998-2008. A study of these four information science journals references would identify which disciplines were cited by the information science journals and could be very helpful in understanding the relationship between subject areas of information science and other subject disciplines and further to discover the knowledge flowing into this field. The results of this study may help seeking the answer for the following questions:

(1) What core journals have been cited by information science journals?

(2) What main class and sub-class are for the cited journals?

(3) What subjects are for the cited journals for information science?

(4) What books have been cited most by information science journals?

(5) What main class, sub-class and subject are for the cited books?

The answers of these questions may clarify the knowledge input for the domain of information science.

2. Methods

There are two assumptions for this study:

(1) *JASIST*, *IPM*, *JOD* and *JIS* accurately represents the information science discipline. This may be justified as these four journals are leading journals in the subject.

(2) The 11 volumes, from 1998 to 2008, for the studied four journals, examined are a fair sample of *JASIST*, *IPM*, *JOD* and *JIS*. At least it reflects the characteristics of the studied period.

2.1 Data collection

For each volume of the studied journals from electronic journals in the Library of National Chenchi University and Tamkang University in Taiwan, from 1998 to 2008, full-length scholarly papers, including research articles and review articles, plus the brief communications were identified. Other type of materials, such as bibliographies, abstracts sections, book reviews, letters, obituaries, announcements, news items, conference reports, committee reports, features, and editorials were excluded in the analysis. The references of each article on electronic version were downloaded. (Hsu,

2009; Fang, 2009) Only journal and book references were considered. References other than these two types, such as theses or reports, were excluded.

2.2 Data analyses

The data collected were processed using Excel by Hsu (2009) and Fang (2009). All the research and review articles published in the four studied journals in the study period and the nature of references cited are analyzed. The present work focuses on the subject of journal and book references cited in the papers published in the four studied journals from 1998 to 2008. The present study also identifies the amount of journals and books cited and explore the subject matter of these publications.

The subject scope of the core cited journals determined using Bradford Law was identified by Ulrich's International Periodical Directory (2010) database of Ulrichsweb as well as the web page of the journal. This study further identifies the main class and subclass of cited journals from Ulrich's International Periodical Directory (2010) and OCLC WorldCat (2010) on the basis of Library of Congress Classification (LCC). The classification was mainly based on LCC, and supplemented with Dewey Decimal Classification (DDC). In LCC, the first character symbolizes the main class, and second character represents subclass. For example, the journal *D-Lib Magazine* is one of the cited journals, which is categorized in WorldCat with a classification number of ZA 4080. For which Z represents the main class of Bibliography. Library science. Information resources (General), ZA stands for information resources (General), and 4080 represents Digital libraries (Fang, 2009). If journals were classified by DDC, the corresponding LCC number would be examined according to the Dewey-LC Conversion table made by OCLC (2010). If the corresponding LCC number could not be found, the data would not be analyzed.

The subject of cited journals for library and information science, which contributed about 50 percent of cited references, was examined on the basis of the descriptor field of each record in the Library and Information Science Abstracts (LISA) (2010). The descriptor field utilized controlled vocabulary from a thesaurus or from subject headings list that were created by the database producer. As indicated by Lancaster (1986), a controlled vocabulary would control the synonym, nearly synonyms, homographs, and related terms; therefore, the search for a descriptor field would retrieval items with particular and comprehensive subject meanings.

On the other hand, the main class, subclass and the subject of cited books were identified by LCC and Library of Congress Subject Headings (LCSH) searching from OCLC WorldCat.

3. Results

3.1 Cited journals

This study first explores the distribution and subjects of references in four information science journals under study from 1998 to 2008. In total, there are 2913 research articles and review papers published in the four journals during the study period. These four journals in the study period cited 105,063 references. Among the cited references, journal literature is the major reference sources.

3.1.1 Highly cited journals. There are 4,668 journals cited by the four information science journals under study, constituting 52,822 (50.3 percent) cited times. As demonstrated in Table I, the top 30 highly cited journals, which were cited at least 200 times by the four journals under study, constitute 50 percent of all citations. The top five cited journals are *JASIST* (15.4 percent), *IPM* (6.7 percent), *JOD* (4.7 percent),

Rank	Title	Times cited	%	Cumulative times cited	Cumulative %
1	*Journal of The American Society for Information Science and Technology*	8,128	15.39	8,128	15.39
2	*Information Processing and Management*	3,519	6.66	11,647	22.05
3	*Journal of Documentation*	2,499	4.73	14,146	26.78
4	*Scientometrics*	1,736	3.29	15,882	30.07
5	*Journal of Information Science*	1,228	2.32	17,110	32.39
6	*Communications of The ACM*	1,112	2.11	18,222	34.50
7	*Annual Review of Information Science And Technology*	831	1.57	19,053	36.07
8	*Library And Information Science Research*	575	1.09	19,628	37.16
9	*ACM Transactions On Information Systems*	480	0.91	20,108	38.07
10	*Library Quarterly*	460	0.87	20,568	38.94
11	*Library Trends*	437	0.83	21,005	39.77
12	*Science*	435	0.82	21,440	40.59
13	*College & Research Libraries*	415	0.79	21,855	41.37
14	*International Journal Of Human-Computer Studies*	399	0.76	22,254	42.13
15	*D-Lib Magazine*	376	0.71	22,630	42.84
16	*Nature*	356	0.67	22,986	43.52
17	*MIS Quarterly*	319	0.60	23,305	44.12
18	*Aslib Proceedings*	306	0.58	23,611	44.70
19	*Social Studies of Science*	235	0.44	23,846	45.14
20	*Journal of the Medical Library Association*	234	0.44	24,080	45.59
21	*Sigir Forum*	229	0.43	24,309	46.02
22	*Computer Networks*	228	0.43	24,537	46.45
23	*Computational Linguistics*	228	0.43	24,765	46.88
24	*Organization Science: A Journal of The Institute of Management Sciences*	224	0.42	24,989	47.31
25	*Online Information Review*	219	0.41	25,208	47.72
26	*Journal of Academic Librarianship*	208	0.39	25,416	48.12
27	*Reference and User Services Quarterly*	207	0.39	25,623	48.51
28	*Information Retrieval*	207	0.39	25,830	48.90
29	*Information Research*	203	0.38	26,033	49.28
30	*ACM Computing Surveys*	200	0.38	26,233	49.66
31	*Management Science*	193	0.37	26,426	50.03
Other	4,637 titles	26,396	49.97	26,396	49.97
Total	4,668 titles	52,822	100	52,822	100

Table I.
Journals of top 50 percent cumulative cited papers in the four journals under study

Scientometrics (3.3 percent), *JIS* (2.3 percent). The other highly cited journals include *Communications of the ACM*, *Annual Review of Information Science and Technology*, *Library and Information Science Research*, *ACM Transactions on Information Systems*, *Library Quarterly*. *JASIST*, *IPM*, *JOD* and *JIS* are the source journals of this study. They themselves are the four of the top five cited journals contributing about 29 percent of citations. This demonstrates a high tendency of self-citation for these journals.

3.1.2 Bradford distribution and the core of the cited journals. Bradford's law has been widely used to study the journal literature distribution. The distribution of the journal literature cited in the four journals during the study period was fit to Bradford's law by plotting the cumulative number of papers for each journal versus the logarithm of its rank. The plot thus obtained, as shown in Figure 1, is not quite similar to the

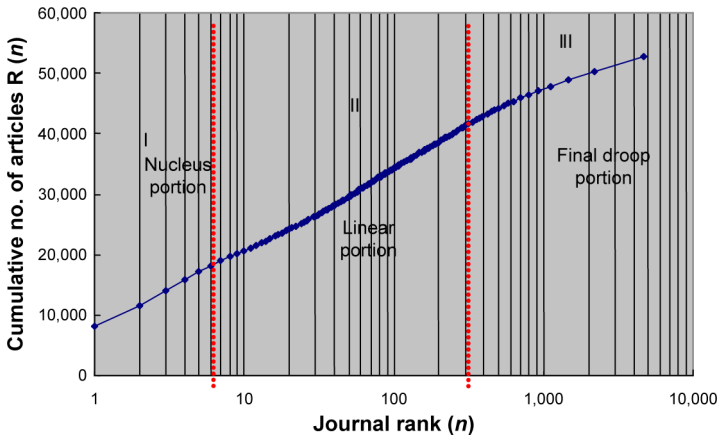

Figure 1.
Bradford-Zipf's distribution of journal literature cited in the four journals under study

typical Bradford-Zipf plot that has a characteristic smooth S-shaped curve, with the final droop portion lying below the linear portion of the curve. The figure indicates that the curve rises gradually and nonlinearly for the first top ten journals, about 40 percent of the total citation, followed by an approximately linear portion, from rank 11 to about 500, after that the curve goes through a droop portion as is typical for the Bradford-Zipf plot. The top ten journals, cited 20,568 times may be considered as the core journals that were cited in the four journals, i.e. in the information science domain. The top 30 journals contain 26,233 papers (50 percent of the 52,822 cited literature); 4,638 journals include the remaining 50 percent of articles. The characteristics of the core cited journals are further investigated in the following section.

The presence of the final droop portion indicates that the cited journal literature in the information science domain has been widely spread to many different journals. Hawkins (1978) suggested that this phenomenon might be due to the dispersion of the cited literature. This is consistent with the fact that 2,387 journals, nearly 50 percent of the cited journals, were cited only once. The scattering of cited journal literature confirms that the information science is a multidisciplinary subject.

3.1.3 Core cited journals. The description of subject scope of the ten core journals identified from the Bradford plot is summarized in Table II. Table II indicates that six out of these ten journals, namely, *Journal of the American Society for Information Science and Technology*, *Information Professing and Management*, *Journal of Documentation*, *Journal of Information Science*, *Annual Review of Information Science and Technology*, *Library and Information Science Research* cover subjects dealing with information science. These are information science oriented journals, which are the primary knowledge flow into information science itself.

The second main stream of knowledge flow into the information science is computer science. Table II shows that two ACM journals, *Communications of the ACM* and *ACM Transactions on Information Systems*, are computer science oriented. Table II also illustrates that *Library Quarterly* and *Scientometrics* are also in the core cited journal list indicating the information science also receives knowledge from library science and quantitative study of science. Indeed, *Library Quarterly* is a traditional library science journal and *Scientometrics* is a journal focusing on quantitative aspects of the science communication and science policy.

Rank	Journal title	Times cited	Subject scope
1	*Journal of The American Society for Information Science and Technology*	8,128	Theory of information; communication; management, economics and marketing; applied information science; social and legal aspects of information
2	*Information Processing and Management*	3,519	Theory, principles, and procedures in information processing; processes of communication among humans and between humans and machines; modeling and evaluation of information system performance; management and economics of information and information systems; information policies
3	*Journal of Documentation*	2,499	Information science, librarianship; information and knowledge management; information and knowledge organization; information seeking and retrieval, and human information behavior; information and digital literacies
4	*Scientometrics*	1,736	Quantitative aspects of the science of science, communication in science and science policy
5	*Journal of Information Science*	1,228	Information science theory, policy, application or practice; information creation, organization, storage, communication and utilization of information and knowledge resources
6	*Communications of The ACM*	1,112	Emerging areas of computer science, new trends in information technology, and practical applications
7	*Annual Review of Information Science And Technology*	831	New trends and significant developments on information science and technology
8	*Library and Information Science Research*	575	Research process and research findings in library and information science
9	*ACM Transactions On Information Systems*	480	Design and evaluation of computer software
10	*Library Quarterly*	480	Historical, sociological, cultural, evaluative, statistical, bibliographical, managerial and educational of librarianship

Table II.
Subject scope of core cited journals for the four journals under study

3.1.4 Main class and subclass of cited journals. Totally there are 20 main classes of journals cited in four journals as shown in Table III. Among them "bibliography. library science. information resources (general)" is the most dominant one contributing nearly half (49.6 percent). The second and third highly cited journals were that categorized under

Rank	Main class	Times cited	Percent	Cumulative %
1	Z – Bibliography. Library Science. Information Resources (General)	25,506	49.56	49.56
2	Q – Science	10,289	19.99	69.55
3	H – Social Sciences	6,100	11.85	81.40
4	T – Technology	3,020	5.87	87.26
5	R – Medicine	1,697	3.30	90.56
6	B – Philosophy. Psychology. Religion	1,419	2.76	93.32
7	P – Language and Literature	1,193	2.32	95.64
8	L – Education	1,106	2.15	97.79
9	K – Law	280	0.54	98.33
10	J – Political Science	266	0.52	98.85
11	A – General Works	158	0.31	99.15
12	G – Geography. Anthropology. Recreation	154	0.30	99.45
12	C – Auxiliary Sciences of History	96	0.19	99.64
14	N – Fine Arts	53	0.10	99.74
15	D – World History and History of Europe, Asia, Africa, Australia, New Zealand, etc.	40	0.08	99.82
16	E/F – History of the Americas	31	0.06	99.88
17	M – Music and Books on Music	24	0.05	99.93
17	U – Military Science	21	0.04	99.97
19	S – Agriculture	16	0.03	100.00
20	V – Naval Science	1	0.00	100.00
Total		51,470[a]	100	

Notes: [a] Some of journals may not be found in the *Ulrich's Periodical Directory*; consequently, the total cited times of main class is less than the total cited times, i.e. 52,822, of journals cited by the four studied journals in the study period

Table III.
Main class of journals cited in the four journals under study

"science" (20.0 percent) and social sciences (11.9 percent). This table suggests that library science is the most cited class of journals, and followed by science and social sciences.

On the other hand, there were 141 subclasses of journals cited in the papers of the four studied journals. Table IV presents the top ten subclasses. Again, the most dominant subclass is library science related: "books(general). writing. paleography. book industries and trade. libraries. bibliography" (48.1 percent), followed by "electronic computers. computer science" (9.1 percent). It should be noted that highly cited journals, e.g. *JASIST*, *JOD* are classified in the general bibliography class. Generally speaking, the result is consistent with that of main classes. Papers published in the four journals mainly cited journals dealing with general bibliography, libraries, book industries and trade, and computer science during the study period.

3.1.5 Subjects of cited journals of library and information science. By examining the descriptor field of each record in the Library and Information Science Abstract (LISA) database, Table V illustrates the percentage, in descending order, of cited frequency for each subject term of library and information science papers cited by the four journals of this study. There were 4,305 unique subject terms cited 104,590 times in the LIS articles. The most cited subject was "searching" (4.99 percent), followed by "online information retrieval" (3.99 percent). "Information work", "subject indexing", "information storage and retrieval", "World Wide Web" come next in order.

Rank	Subclass	Times cited	Percent
1	Books (General). Writing. Paleography. Book industries and trade. Libraries. Bibliography	24,751	48.09
2	Electronic computers. Computer science	4,662	9.06
3	Science (General)	4,259	8.27
4	Industries. Land use. Labor	2,261	4.39
5	Commerce	1,817	3.53
6	Psychology	1,293	2.51
7	Electrical engineering. Electronics. Nuclear engineering	1,108	2.15
8	Technology (General)	1,059	2.06
9	Medicine (General)	967	1.88
10	Philology. Linguistics	833	1.62
Percentage of top ten subclasses			83.56
Percentage of other subclasses			16.44
Total percentage			100
Kinds of subclasses		141	

Table IV.
Top ten subclasses of journals cited in the four journals under study

Table V suggests that the information science papers tend to deal with issues related to information retrieval, information seeking behavior, citation analysis, use and user study, search strategy, world wide web, etc. The subject USA and UK perhaps mean that articles about USA or UK organizations, such as Thelwall's article in 2002, "The top 100 linked pages on UK university Web sites: high backlink counts are not usually directly associated with quality scholarly content."

3.2 Analysis of cited books

There are 10,654 titles of book cited for 21,212 times by the four information science journals for the 11 selected years under study. On average, every title was cited 1.4 times. All these book references can be divided into 19 main classes, 154 subclasses and 5,948 subjects.

3.2.1 Highly cited books. Table VI lists the ten most highly cited books by the four journals for the time period under study. *The Introduction to Modern Information Retrieval*, authored by Salto, G. and McGill, M., is the top most cited book with total cited times of 227. The number two and number three highly cited books are *The SMART Retrieval System: Experiments in Automatic Document Processing*, authored by Salton, G. and *Information Retrieval*, authored by Van Rijsbergen, C.J. Interestingly, Salton, G. and Van Rijsbergen, C.J. are also among the top most cited monographic book authors in DeHart (1992) study on "monographic references and information science journal literature" for the period of 1987-1990. They remain to be among the most cited monographic authors and their cited times is significantly increased. In DeHart work, they were cited for 40 times and 21 times, respectively, in four years, which is much less than the current 144 to 227 times in 11 years.

Among the ten most cited books, G. Salton authored two books and co-authored the top most cited books. This suggests Salton is the most influencing book writer for the authors in the four key journals under study. With regards to the subjects of the book cited, the keywords in the book title of the ten most cited ones suggest there are eight books related to information retrieval or information seeking. This demonstrates that information retrieval and information seeking are the two-most important subjects for the books cited in the four journals under study. Indeed, this is the case, as will be

Rank	Subject (descriptor)	Times cited	Percent	Cumulative %
1	Searching	5,221	4.99	4.99
2	Online information retrieval	4,177	3.99	8.99
3	Information work	3,753	3.59	12.57
4	Subject indexing	2,701	2.58	15.16
5	Information storage and retrieval	2,675	2.56	17.71
6	World wide web	2,498	2.39	20.10
7	Technical services	2,226	2.13	22.23
8	Computerized information storage and retrieval	1,604	1.53	23.76
9	Citation analysis	1,541	1.47	25.24
10	Information seeking behavior	1,501	1.44	26.67
11	Research	1,411	1.35	28.02
12	Evaluation	1,368	1.31	29.33
13	Bibliometrics	1,269	1.21	30.54
14	Computerized information retrieval	1,213	1.16	31.70
15	Periodicals	1,175	1.12	32.83
16	Library materials	1,165	1.11	33.94
17	Internet	1,036	0.99	34.93
18	Relevance	1,007	0.96	35.89
19	Information science	895	0.86	36.75
20	Models	855	0.82	37.57
21	User surveys	788	0.75	38.32
22	Articles	734	0.70	39.02
23	Strategies	685	0.65	39.68
24	Web sites	655	0.63	40.30
25	User services	651	0.62	40.93
26	Search strategies	650	0.62	41.55
27	Services	634	0.61	42.15
28	Use	616	0.59	42.74
29	Search engines	614	0.59	43.33
30	USA	612	0.59	43.91
31	University libraries	601	0.57	44.49
32	Computerized subject indexing	582	0.56	45.05
33	User behavior	561	0.54	45.58
34	Students	553	0.53	46.11
35	User needs	519	0.50	46.61
36	Libraries	499	0.48	47.08
37	Electronic media	482	0.46	47.54
38	UK	473	0.45	48.00
39	Performance measures	461	0.44	48.44
40	Librarianship	453	0.43	48.87
41	Surveys	450	0.43	49.30
42	Medicine	443	0.42	49.72
43	Information communication	432	0.41	50.14
44	Online databases	416	0.40	50.54
45	Science	400	0.38	50.92
46	Mathematical models	393	0.38	51.29
47	Users	390	0.37	51.67
48	Links	390	0.37	52.04
49	Theories	378	0.36	52.40
50	Retrieval performance measures	376	0.36	52.76
Other subjects		49,408	47.24	100
Total		104,590	100	

Table V.
Top 50 subjects of journal papers cited in the four journals under study

Rank	Title	Author	Times cited
1	*Introduction to Modern Information Retrieval*	Salton, G. and McGill, M.J.	227
2	*The SMART Retrieval System: Experiments in Automatic Document Processing*	Salton, G.	163
3	*Information Retrieval*	Van Rijsbergen, C.J.	144
4	*Information Retrieval: Data Structures & Algorithms*	Frakes, W.B. and Baeza-Yates, R.	132
5	*Seeking Meaning: a Process Approach to Library and Information Services*	Kuhlthau, C.C.	117
6	*Modern Information Retrieval*	Baeza-Yates, R. and Ribeiro-Neto, B.	115
7	*Automatic Text Processing: the Transformation, Analysis, and Retrieval of Information by Computer*	Salton, G.	112
8	*Information Seeking in Electronic Environments*	Marchionini, G.	92
9	*The Web of Knowledge: a Festschrift in Honor of Eugene Garfield*	Cronin, B. and H.B. Atkins	91
10	*Encyclopedia of library and information science*		86

Table VI.
Top ten highly cited books in the four journals under study

discussed later in section 4.7. The present analysis also reveals that about 90 percent of the books cited 1 to 3 times only. This suggests that the knowledge input from the books cited in these four leading journals, i.e. of the domain of information science, spreads quite extensively.

3.2.2 Main classes and subclasses of cited books. Based on the Library of Congress Classification (LCC), all books that were cited by the four information science journals were grouped into 19 main classes as shown in Table VII. It can be seen from Table VII that "science" (24.3 percent) is the most cited class, followed by "bibliography. library science. information resources (general)" (21.7 percent), "social sciences" (20.1 percent), "philosophy. psychology. religion" (9.6 percent), and "language and literature" (8.4 percent).

Table VIII demonstrates the top ten subclasses of books cited in papers of the four journals. Among the 154 kinds of subclasses of books cited in the four journals, the most cited subclass is "books (general). writing. paleography. book industries and trade. libraries. bibliography" (19.6 percent), followed by "electronic computers. computer science" (15.6 percent), and other subclasses about general science, philology, linguistics, industries, land use and labor, and psychology with contribution greater than 5 percent. The top three most cited and 6th subclasses are exactly the same as the top three most cited and 6th subclasses of journals.

3.2.3 Subjects of cited books. Through the retrieval from the WorldCat, 10,654 books cited by the four information science journals, containing 5,948 unique subject headings with total 43,617 cited times. Subjects cited once account for 63.5 percent (2,902 kinds). This shows the degree of spreading of cited subjects, consistent with that for cited journals. Table IX displays top 30 subjects of books cited on the four information science journals under study.

The top most cited subjects, at least contributing 0.4 percent of total citations, accounts for 19.54 percent of the total citation, even less than one quarter of the total

Rank	Main class	Times cited	%	Cumulative %
1	Q – Science	4,610	24.34	24.34
2	Z – Bibliography. Library science. Information resources (general)	4,104	21.67	46.01
3	H – Social sciences	3,808	20.11	66.11
4	B – Philosophy. Psychology. Religion	1,826	9.64	75.76
5	P – Language and literature	1,590	8.39	84.15
6	T – Technology	1,020	5.39	89.54
7	L – Education	649	3.43	92.96
8	R – Medicine	233	1.23	94.19
9	K – Law	202	1.07	95.26
10	J – Political science	192	1.01	96.27
11	G – Geography. Anthropology. Recreation	176	0.93	97.20
12	N – Fine arts	120	0.63	97.84
13	D – World history and history of Europe, Asia, Africa, Australia, New Zealand, etc.	115	0.61	98.44
14	A – General works	106	0.56	99.00
15	M – Music and books on music	49	0.26	99.26
16	C – Auxiliary sciences of history	48	0.25	99.51
17	U – Military science	47	0.25	99.76
18	E/F – History of the Americas	38	0.20	99.96
19	S – Agriculture	7	0.04	100.00
	Total	18,940	100.00	

Table VII.
Main classes of books cited in the four journals under study

Rank	Subclass	Times cited	%	Cumulative %
1	Books (General). Writing. Paleography. Book industries and trade. Libraries. Bibliography	3,710	19.59	19.59
2	Electronic computers. Computer science	2,963	15.64	35.23
3	Science (General)	1,288	6.80	42.03
4	Philology. Linguistics	1,159	6.12	48.15
5	Industries. Land use. Labor	1,157	6.11	54.26
6	Psychology	1,130	5.97	60.23
7	Sociology (General)	966	5.10	65.33
8	Theory and practice of education	507	2.68	68.00
9	Social sciences (General)	445	2.35	70.35
10	Electrical engineering. Electronics. Nuclear engineering	427	2.25	72.61

Table VIII.
Top ten subclasses of books cited in the four journals under study

citation. This reveals that the subjects of the books cited by four key journals of information science are quiet dispersed. In other words, the authors of these leading journals concern about many related issues. Table IX also indicates that the top two most cited subjects of books cited are information storage and retrieval systems and information retrieval, accounting for 4.07 percent, which is consistent with the top most cited books shown earlier in Table VI. Computer science related, such as human-computer interaction (ranked 4th), artificial intelligence (ranked 6th), database management (ranked 10th), text processing (computer science) (ranked 11th), computer algorithms (ranked 13th), user interfaces (computer system) (ranked

Rank	Subject	Times cited	%	Cumulative %
1	Information storage and retrieval systems	1,237	2.84	2.84
2	Information retrieval	538	1.23	4.07
3	Information science	498	1.14	5.21
4	Human-computer interaction	433	0.99	6.20
5	Indexing	327	0.75	6.95
6	Artificial intelligence	282	0.65	7.60
7	Cognition	277	0.64	8.24
8	Information behavior	266	0.61	8.85
9	Science – philosophy	260	0.60	9.44
10	Database management	253	0.58	10.02
11	Text processing (computer science)	251	0.58	10.60
12	Social sciences – research	243	0.56	11.15
13	Library science	236	0.54	11.69
14	Computer algorithms	234	0.54	12.23
15	Knowledge, theory of	231	0.53	12.76
15	Reference services (libraries)	220	0.50	13.27
17	Knowledge management	218	0.50	13.77
18	Science – social aspects	216	0.50	14.26
19	Social sciences – statistical methods	210	0.48	14.74
20	Information technology – social aspects	207	0.47	15.22
21	Communication in science	206	0.47	15.69
22	Information technology	203	0.47	16.15
23	Information society	197	0.45	16.61
24	User interfaces (computer systems)	194	0.44	17.05
25	Information resources management	193	0.44	17.49
26	Libraries	188	0.43	17.92
26	Organizational learning	179	0.41	18.33
28	Communication	177	0.41	18.74
28	Library research	175	0.40	19.14
30	Computational linguistics	175	0.40	19.54
Other subjects		3,5093	80.46	100
Total		43,617	100	

Table IX.
Top 30 subjects of books cited in the four journals under study

24th) and computational linguistics (ranked 30th) is the other significant subject of the cited books and accounts for about 4.18 percent of total citations. Library related directly is the third most cited subjects. As shown in Table IX, the total citation for the library science (ranked 13th), reference services (libraries) (ranked 16th), libraries (ranked 26th), and library research (ranked 28th) are cited 819 times, accounting for 1.88 percent.

The authors of these four leading journals also pay significant attention on the social science related subjects such as social sciences research (ranked 12th), science-social aspects (ranked 16th), social science-statistical methods (ranked 19th), and information technology-social aspects (ranked 20th). As pointed out earlier, this demonstrates the subject evolution toward social science as times evolves.

3.3 Summary

The significant results of the present study are summarized in the following:

- There are 4,668 journals cited by the four information science journals under study, constituting 52,822 (50.3 percent) cited times. The top five cited journals

are *JASIST* (15.4 percent), *IPM* (6.7 percent), *JOD* (4.7 percent), *Scientometrics* (3.3 percent), *JIS* (2.3 percent). The majority of the core cited journals are information science oriented.

- Totally there are 20 main classes of journals cited in four journals. Among them "bibliography. library science. information resources (general)" is the most dominant one contributing nearly half (49.6 percent). The second and third highly cited journals are that categorized under "science" and "social sciences".
- The highly cited subjects of library and information science journals encompass searching, online information retrieval, information work, subject indexing, information storage and retrieval, world wide web, technical services, computerized information storage and retrieval, citation analysis, information seeking behavior, etc.
- The most cited subclass of cited journals is "books(general). writing. paleography. book industries and trade. libraries. bibliography", which accounts for 48 percent.
- The top three most cited books in order are: *The Introduction to Modern Information Retrieval*, authored by Salto, G. and McGill, M., *The SMART Retrieval System: Experiments in Automatic Document Processing*, authored by Salton, G. and *Information Retrieval*, authored by Van Rijsbergen, C.J.
- For the books cited, "science" is the most cited class, followed by "bibliography. library science. information resources (general)" and "social sciences"; while the most cited subclass is "books (general). writing. paleography. book industries and trade. libraries. bibliography", followed by "electronic computers. computer science".
- Subjects of books highly cited in the four journals under study were about information retrieval system, information science, human-computer interaction, indexing, information behavior, social sciences research, library science and different kinds of libraries.

4. Discussion

4.1 Discussion on key findings

Excluding *Scientometrics*, four of the top five cited journals, namely *JASIST*, *IPM*, *JOD* and *JIS*, are information science oriented journals, as Paisley (1990) and McCarthy (2000) identified in their studies on the multidisciplinary nature and time evolution of information science. These four journals are the source journals of this study. This implies that the domain of information science possessing strong self knowledge flow.

The study of the core cited journals indicated that the majority of the core cited journals are information science oriented. This is another phase of "self-citation". The researchers in the information science tend to cite more research outcomes in their own subject fields. This is probably also true for other disciplines as researchers are generally more familiar with their own subject fields. The subjects of the core cited journals demonstrate the main stream of knowledge flow into information science is itself.

The top main class of the cited journals in papers published in the four information science journals is bibliography. library science. general information resources which contributes 50 percent, while the most cited subclass of cited journals is "books(general). writing. paleography. book industries and trade. libraries. bibliography", which accounts for 48 percent. On the other hand, the most cited subclass of the cited books is the same as the cited journals with but the

percentage is much low value of 20 percent. This suggests library science, book industries and trade, general information resources consist of a mainstream of knowledge flow into information science on the basis of the LCC. This result is consistent with DeHart's (1992) study that books of generalities of the Dewey main class (000) are cited most.

However, the second and the third highly cited journals were "science" (20.0 percent) and "social sciences" (12 percent), respectively. For the highly cited books, "science" (24.3 percent) is the most cited class, followed by "bibliography. library science. information resources (general)" (21.7 percent), "social sciences" (20.1 percent), "philosophy. psychology. religion" (9.6 percent), and "language and literature" (8.4 percent).This suggests significant knowledge flow into information science from science or social science subjects.

As library and information science is the main class contributing about 50 percent journal literature on the knowledge flow into the information science, further analysis indicates that the highly cited subjects of library and information science journals encompass searching, online information retrieval, information work, subject indexing, information storage and retrieval, world wide web, technical services, computerized information storage and retrieval, citation analysis, information seeking behavior, etc. On the other hand, the most cited LCSH subjects on WorldCat of books are information storage and retrieval systems, information retrieval, information science, human-computer interaction etc. Astrom (2007) also observed the dominance of web-related studies in the field of library and information science.

Comparing to earlier studies, the subjects of the highly cited books has spread to other subjects. The highly cited books in the four journals under study cover subjects such as information retrieval system, information science, human-computer interaction, indexing, information behavior, social sciences research, library science and different kinds of libraries. DeHart (1992) also identified the five subjects, supplied from OCLC records, appearing most often were information storage and retrieval systems; artificial intelligence; discourse analysis; database management; and human-computer interaction. Comparison with DeHart study in 1992, one can see the subject evolution with time of cited book in the area of information science. Information retrieval system and human-computer interaction remain most cited, while the cited books spread to other non-information technology subjects, such as information behavior and social science research.

4.2 Limitations

The two assumptions made earlier in the section of Methods also put possible limitations of the present study that the findings are reflected by the four leading journals in the field of information science in the selected study period.

Besides, as pointed out in the section of Data Collection, only journals and book references cited by the research and review articles in the four selected journals are considered in the present study. These two types of references cited account for about 80 percent. This may impose a limitation on the present study, though it is insignificant. In addition, LISA descriptor has been employed for the identification of the subject of the journal articles in the field of library and information science. However, some of new research developments, e.g. webometrics, web usability, web use, web impact factor, web link, informetrics, mapping and research performance were not displayed in the LISA descriptor field. The reason may be due to that the subject scope of LISA descriptor is too broad to cover the related term of WWW, internet and informetrics.

4.3 Implications

The results of the present study demonstrate how the domain knowledge of information science had been influenced mainly by itself and partially by other disciplines such as sciences and social sciences. The findings of this work should have great interest for citation study in library and information science for understanding the nature of information science discipline.

5. Conclusion and further studies

The present study conducts a bibliometric and citation analysis on the cited journal and book literature of four information science journals for volumes published in 11 selected years. The results of the present study reveal that the knowledge input for information science field has three basic themes: experimental information retrieval focusing on user-system interaction, user and its behavior and IR system development. This study also found Web has had an effect on the knowledge structure of information science. New subjects such as world wide web, internet, web sites helped enhancing information science research.

Moreover, information science, as represented by the four information science journals, i.e. *JASIST*, *IPM*, *JOD* and *JIS*, is found to be a developing interdisciplinary subject with an expanding citing literature. Increasingly, there has been great growth in the citing of previous literature in "information science" as well as "social science" and "science" papers. The knowledge inputs for information science include mainly information science itself and social sciences and general science as well. In addition, there are minor inputs from various subjects.

The present study investigates the knowledge flow into the information science. It also will be very interesting to study the knowledge flow out of the domain of information science and to do a comparison between these two types of knowledge flow of the field.

References

Astrom, F. (2007), "Changes in the LIS research front: time-sliced co-citation analysis of LIS journal articles, 1990-2004", *Journal of the American Society for Information Science and Technology*, Vol. 58 No. 7, pp. 947-57.

Bonnevie, E. (2003), "A multifaceted portrait of a library and information science journal: the case of the Journal of Information Science", *Journal of Information Science*, Vol. 29 No. 1, pp. 11-23.

Borgman, C.L. (1999), "Books, bytes, and behavior: rethinking scholarly communication for a global information infrastructure", *Information Services & Use*, Vol. 19 No. 2, pp. 117-21.

Borko, H.D. (1968), "Information science: what is it?", *American Documentation*, Vol. 19 No. 1, pp. 3-5.

DeHart, F.E. (1992), "Monographic references and information science journal literature", *Information Processing & Management*, Vol. 28 No. 5, pp. 629-35.

Desai, C.M. (2003), "Getting cited: ten tips for practitioners of citation analysis in the library", *College and Research Libraries News*, Vol. 64 No. 1, p. 21.

Dewey-LC Conversion table made by OCLC (2010), available at: www.questionpoint.org/crs/html/help/es/ask/ask_map_ddctolcc.html

Fang, B.-L. (2009), "Citation analysis of Western journal literature on information science: a case study of JASIS(T)", Master Thesis, Department of Information and Library Science, Tamkang University, New Taipei City, Taiwan.

Harter, S.P. (1996), "The impact of electronic journals on scholarly communication: a citation analysis", *Public Access Computer Systems Review*, Vol. 7 No. 5, pp. 5-34.

Hawkins, D.T. (1978), "Bibliometrics of the online information retrieval literature", *Online Review*, Vol. 2 No. 4, pp. 345-51.

Hsu, J.-Y. (2009), "Citation analysis of Western journal literature on information science", Master thesis, Graduate Institute of Library, Information and Archival Studies, National Chengchi University, Taipei City, Taiwan.

Lancaster, F.W. (1986), *Vocabulary Control for Information Retrieval*, 2nd ed., Information Resources, Arlington, VA.

Library and Information Science Abstracts (LISA) (CSA) (2010), available at: http://search.proquest.com/lisashell?accountid=10067

McCarthy, C.A. (2000), "Journal of the century in library and information science", *The Serials Librarian*, Vol. 39 No. 2, pp. 121-38.

Nebelong-Bonnevie, E. and Frandsen, T.F. (2006), "Journal citation identity and journal citation image: a portrait of the Journal of Documentation", *Journal of Documentation*, Vol. 62 No. 1, pp. 30-57.

OCLC WorldCat (2010), available at: www.oclc.org/worldcat/

Paisley, W. (1990), "Information science as a multidiscipline", in Pemberton, J.M. and Prentice, A.E. (Eds), *Information Science: The Interdisciplinary Context*, Neal-Schuman, New York, NY, pp. 3-24.

Persson, O. (1994), "The intellectual base and research fronts of JASIS 1986-1990", *Journal of the American Society for Information Science*, Vol. 45 No. 1, pp. 31-8.

Saracevic, T. (1999), "Information science", *Journal of the American Society for Information Science*, Vol. 50 No. 12, pp. 1051-63.

Ulrich's International Periodical Directory (2010), New York, New York Public Library, available at: https://ulrichsweb.serialssolutions.com/

White, H.D. and McCain, K.W. (1998), "Visualizing a discipline: an author co-citation analysis of information science, 1972-1995", *Journal of the American Society for Information Science*, Vol. 49 No. 4, pp. 327-55.

Zhao, D. and Strotmann, A. (2008), "Information science during the first decade of the web: an enriched author cocitation analysis", *Journal of the American Society for Information Science and Technology*, Vol. 59 No. 6, pp. 916-37.

About the author

Dr Ming-yueh Tsay is a Professor and Chairperson of the Graduate Institute of Library, Information and Archival Studies at National Chengchi University (NCU), Taipei, Taiwan. Before joining NCU in August 2004, she had been a faculty member of the Department of Information and Library Science at Tamkang University, Taiwan for 15 years serving as Associate Professor and Professor. She served as the Head of the Department from August 2003 to July 2004. She was an Associate Researcher in the libraries of Industry Technology Research Institutes from July 1986 to July 1989. She earned her MS, CAS and PhD degrees in the Graduate School of Library and Information Science from University of Illinois at Urbana-Champaign in 1983, 1986 and 1996, respectively. Her research activities for the past 20 years have been in the areas of bibliometrics, information retrieval, library and information science. Ming-yueh Tsay can be contacted at: mytsay@nccu.edu.tw