A Comparative Analysis on Keywords of International and Korean Journals in Library and Information Science

국내외 문헌정보학 저널의 키워드 비교 분석

Eungi Kim*

초 록

The aim of this study was to discover various Library and Information Science (LIS) research areas by examining similarities and differences between LIS journals in terms of keyword characteristics. To conduct this study, for the years from 2004 to 2016, the keywords of 6 international journals were downloaded from Scopus database (http://www.scopus.com), and the keywords of 4 Korean journals were downloaded from the RISS database (http://www.riss.co.kr). The characteristics of keywords were investigated by examining frequently used keywords and frequently used distinctive keywords pertaining to international and Korean journals. The distinctive keywords are referred to as the keywords that appear in one domain but not in another. The result of this study indicated the following: a) a frequency analysis of the keywords showed major research themes and unique traits concerning Korea. b) In general, the keywords used in Korean journals frequently reflected the library as a major subject area of research, while keywords used in international journals reflected bibliometrics and information retrieval as major subject areas of research. c) The overarching themes of each created dataset were clearly noticeable in frequently used distinctive keywords. d) Some keywords were bound by a nation or by a region due to their scope of usage. The important implication of this study is that both most frequently used keywords and most frequently used distinctive keywords seemed to adequately represent the LIS subject areas

Keywords: Keywords, Frequently used keywords, Distinctive keywords, Library and Information Sciences, Journal, Subject areas, Research areas

ABSTRACT

본 연구의 목적은 키워드 특징 면에서 문헌정보 저널에서 나타나는 유사점과 차이점을 조사하여 다양한 문헌 정보학 연구 영역을 발견하는 데 있다. 이 연구를 수행하기 위해 2004 년부터 2016 년까지 네 개의 한국 저널의 키워드가 RISS 데이타베이스에서 수집 되었고(http://www.riss.co.kr) 그리고 여섯 개의 국제저널의 키워드가 SCOPUS 데이타베이스에서 수집 되었다(http://www.scopus.com). 키워드의 특징은 한국 및 국제저널에 관하여서 자주 사용 되었던 키워드와 자주 사용되었던 독특한 키워드를 검증하는 연구이었다. 독특한 키워드란 한 분야에서는 나타나지만 다른 분야에서는　나타나지 않는 키워드를 말한다. 이 연구의 결과는 다음과 같다. 가) 키워드 빈도 분석 결과는 한국의 문헌정보 학의 연구주제와 연구특색을 보여 주는 것으로 나타났다. 나) 일반적으로 한국 저널에서 사용 된 키워드는 도서관과 관련된 주제의 영역을 나타냈고, 국제 저널에 사용되는 키워드는 서지 측정법과 관련된 주제 영역을 나타냈다. 다) 빈번히 사용되었던 독특한 키워드에서도 이러한 전반적인 연구 테마를 명백히 나타냈다. 라) 어떤 키워드는 쓰이는 범위가 한 국가나 지역으로 한정되어 있는 것으로 나타냈다. 이 연구의 중요한 시사점은 가장 자주 사용되는 키워드와 가장 자주 사용되는 독특한 키워드는 둘 다 문헌정보 학의 주제 영역을 적절하게 반영하고 있는 것으로 보인다는 것이다.

키워드: 키워드, 자주 사용된 키워드, 독특한 키워드, 문헌정보학 연구, 논문지, 저널, 연구 주제, 연구 분야, 연구 영역

Ⅰ. Introduction

Library and Information Science (LIS) is regarded as a meta-discipline that embraces a wide variety of applicable theories, philosophies, and research methods (Bawden & Robinson 2015). Within the disciplines closely related to LIS, a variety of subject areas have been developed. Typical LIS journal focuses on some aspects of LIS, and their aims and scope of these journals are reflected in the subject areas of LIS. Due to the multi-disciplinary and evolving tendencies of the LIS field, it is difficult to identify the extent of subject areas that the LIS journal touches upon.

Meanwhile, practitioners and academics in the LIS community have a long history of recognizing the keywords as valuable metadata that serves multiple functions. Keywords can not only increase the searchability of documents, but they can also be used to enhance readability of the document. As the keywords serve as surrogate information for the readers, the keywords tend to represent some relevant subject matter.

A large number of keywords allows us to reveal some characteristics of published journal articles. Because we can assume that keywords are intended to represent the documents as accurately as possible (Grant 2010), keywords can be utilized in order to discover insightful information regarding a particular LIS journal and LIS research areas pertaining to a particular domain.

Previous works attempted to discover some major themes and subject areas in LIS journals. To extend such an effort, the aim of this study was to discover various aspects of LIS research areas by examining the similarities and differences between in international and Korean journals in terms of keyword characteristics. In particular, frequently used keywords and distinctive keywords were compared among the international and Korean LIS journals. To this end, this study identified similar and dissimilar pattern of themes among journals and between international and Korean LIS journals. As whole, this study will shed some lights on some international and Korean LIS research areas through the keywords of LIS journal articles.

Ⅱ. Related Works

Previous studies have examined keywords of research articles for a variety of purposes. In particular, the characteristics of keywords have been studied from the perspective of increasing search retrieval efficiency. For example, Névéol, Dogan, and Lu (2010) attempted to analyze the keywords for the purpose of improving retrieval of documents. They showed that over 60% of the keywords could be linked to a closely related indexing term in *Medline*. This means that remaining 40% of the keywords might not have adequately represented the contents of their journal articles. Still, authors suggested that using the keywords should be useful for developing medical terminologies.

The keywords of academic publications also have been scrutinized in the anticipation of discovering meaningful characteristics concerning various fields of study. For instance, Barthel and Seidl (2017) attempted to map interdisciplinary collaboration between natural science and social science through the use of keywords and titles using the *Scopus* database. Cunningham and Kwankkel (2011) observed the evolution of the top ten keywords across three field based corpus: engineering management, technology management and management of technology. The authors showed that keywords have changed over time since 1950, as shown by the keywords displayed in the subject matters of concern. Zhang and Hong (2014) used a networked model to describe the intellectual structure related to reading (e.g., reading education). In the process, the authors showed some trends of keyword usage.

Although all of the above mentioned works analyzed keyword usages showed some interesting results, the analysis of keywords have limited applicability to the field of LIS since the keywords may are likely to be disciplinary dependent and keywords characteristics in one field may differ from another field. For example, keywords in medical journals tend to be more precise than those in LIS journals. Regardless, it is difficult to generalize their findings to Korea, since keyword usages in Korean domestic journals may be different from the keywords used in international journals.

Unlike the above mentioned previous works, keywords have been used to uncover various aspects of LIS research by Korean researchers. Lee (2016) analyzed the keywords of LIS journal articles, focusing on the subject of public libraries. In their approach, the characteristics of

keywords were analyzed to find branches of public library research. Lee also concluded that studies on the Korean public library mainly focus on the following areas: management issues, cooperation between public libraries and other kinds of libraries, special users, human resources, small libraries, and lifelong education. Seo et al. (2015) identified notable academic characteristics in regards to the *Journal of the Korean Biblia Society for Library and Information Science* (*JKBSLIS*). Their research was based on 300 articles from 2010 to 2014, and keywords were used to identify main themes of research. According to the authors, the main research areas of JKBSLIS are the following "public library and reading", "academic library and collection management", and "school library and information literacy education". However, not all journals that fall in the category of LIS have been examined by these authors.

Pertaining to LIS research as a whole, other metadata, such as the title field, was utilized to discover the tendencies of LIS research by some researchers. For example, using the title of research articles published between 1988 and 2007, Milojević et al. (2011) identified various research areas that could be categorized under the following major areas: library science, bibliometrics and scientometrics, and information science.

In essence, the previous research suggests that the keywords can be used to identify various characteristics of a domain, including the fields of study, subject areas, country or region, etc. However, there has been a lack of prior research that compared the keywords of journal articles in an attempt to understand the research areas of Library and Information Science. From this perspective, this paper intends to critically analyze keywords pertaining to LIS journals published in Korea and in international journals. By identifying common and less common used keywords, this study characterizes Korean LIS research as a whole.

Ⅲ. Methodology

The main approach employed in this study was to compare the keywords among the Korean LIS journals and the international LIS journals. Since most journals require authors to submit a set of keywords along with the manuscript for publication, keywords nowadays are readily available through bibliographic databases. To conduct this study, 6 international journals were randomly selected using the top 30 *SCImago Journal Rank* (SJR) in the category of LIS. Based

on *Scopus* database (http://www.scopus.com), SJR measures the journal's influence and can be considered as a quality indicator. In this study, an assumption was made that the selected sample journals, within reason, are a small representation of quality journals in the field of LIS. Next, 4 domestic journals from Korea were selected since these journals are considered as factual LIS journals. These Korean journals were indexed by the Korean Citation Index (KCI). Thus, a total of 10 journals − 6 international journals and 4 Korean journals − were selected. The number of journals was limited in order to complete the study in a relatively short duration of time. Moreover, limiting the journal size to 10 provided simplicity of sampling and ease of research.

<Tab. 1> Scope and Aims of LIS International and Korean Journals

	Journal Names	Scopes/Aims
International Journals	*Information Processing and Management (IPM)*	The theory, methods, or application in the field of information science. (IPM Home Page, 2017).
	Journal of Academic Librarianship (JAL)	The problems and issues germane to college and university libraries. (JAL Home Page, 2017).
	Journal of the American Society for Information Science and Technology (JASIST)	The production, discovery, recording, storage, representation, retrieval, presentation, manipulation, dissemination, use, and evaluation of information and on the tools and techniques associated with these processes. (JASIST Home Page, 2017).
	Journal of Documentation (JD)	Theories, concepts, models, frameworks and philosophies related to documents and recorded knowledge. (JD Home Page, 2017).
	Journal of Information Science (JIS)	All areas of research in the sciences of information and knowledge management, which includes information science theory, policy, application or practice that will advance the thinking in the field. (JIS Home Page, 2017).
	Scientometrics	Quantitative features and characteristics of science and scientific research; investigations in which the development and mechanism of science are studied by statistical mathematical methods. (Scientometrics Home Page, 2017).
Domestic Journals	*Journal of the Korean Biblia Society For Library And Information Science (JKBSLIS)* 한국비블리아학회지	Contributing works on the development of library and information science by selecting special topics within the field of LIS. (JKBSLIS Home page, 2017).
	Journal of Korean Society for Library and Information Science (JKSLIS) 한국문헌정보학회	Research work that reinforces the foundation of library and information science research, solves problems associated with in the library in the practice, and contributes to the development of library in practice. (JKSLIS Home page, 2017).
	Journal of Korea Society for Information Management (JKSIM) 한국정보관리학회	LIS research work including the areas of information management, information processing, knowledge creation, sharing and management, data and text mining, management sciences, knowledge-based systems, statistical simulations, methodological studies for information management research. (JKLISS Home page, 2017).
	Journal of the Korean Library and Information Science Society (JKLISS) 한국도서관정보학회지	All aspects of libraries and information science, including practical applications. (JKLISS Home Page, 2017).

Keywords of 6 international journals were downloaded from the Scopus, whereas keywords of 4 Korean journals were download from the *RISS* database (http://www.riss.kr). Since some keywords were not available in all journals, it was decided to use a range of journal publication dates. In this study, only keywords of research articles published from 2004 to 2016 were collected. In order to avoid sampling bias, the keywords of each journal were consistently included every year from 2004 to 2016, without omitting any published year.

In Table 1, while three Korean journals uses the term " Library and Information Science" in their journal name, the international journals do not use " Library and Information Science" in the journal name. Moreover, the scopes and the aims of the international journals are narrower compared to the Korean LIS journals. Apart from this, the Korean journals appear to be accommodating a wide range of LIS research topics.

The *UNIX* based tools such as *sed* and *awk* (Dougherty and Robbins 1997) were used to obtain frequently used keywords in each journal, and frequently used keywords in international and Korean datasets. Furthermore, the tools were used to obtain distinctive keywords in each journal, and distinctive keywords concerning international dataset and Korean datasets. The exact details of these types of keywords are described in the subsequent sections of this paper. Microsoft *Excel* was also used to obtain the basic statistical information.

In performing the frequency counts, the extracted terms were only normalized to a certain extent. For instance, all punctuation marks were removed except for hyphenated words. Capitalized letters were normalized to lower case letters except for acronyms. The stemming tools were not used since the measure of accuracy depends on tool. Instead, only the most frequently used keywords were normalized upon checking singular and plural forms.

Ⅲ. Results

The journal names along with basic statistical information regarding the keywords are shown in Table 2. For each journal, a total number of keywords, a total number of journal articles, total number of unique keywords, the ratio of unique keywords, and the average number of keywords per article are shown. Due to the lexical variations of words, a vast number of keywords were found in the datasets. To this end, the frequency distribution of keywords resembled the *Zipf*

distribution (Newman 1992; Li 2005).

1. The Frequency Count of Keywords

As shown in Table 2, the total number of keywords varied substantially among journals due to the differing number of published articles by each journal. The average number of keywords per article ranges from 2.6 to 5.2 keywords for the international journal, whereas the average number of keywords in the Korean journals ranged from 4.5 to 5.0 keywords per article. Thus, more widespread distribution of the average number of keywords was found in the international journals. The number of unique keywords and the ratio of unique keywords are more important. As shown, the number of unique keywords is defined as unique instances of keywords by squeezing multiple duplicate keywords into one instance. Based on the number of unique keywords, the ratio of unique keywords can be obtained by using the following formula:

ratio of unique keywords = # of unique keywords / total # of keywords

There should be a definite correlation between the ratio of unique keywords and the number of subject matters that the keywords represent. More than likely, a lower ratio would indicate the presence of more duplicate keywords in the dataset, whereas a higher ratio of unique keywords would indicates a greater degree of polarization in the keyword uses in the dataset. As shown, the journal having the lowest ratio of unique keywords is the *JD* (0.41). On the other

⟨Tab. 2⟩ Statistical Information of the Keywords

	Journal Names	Total # of Keywords	# of Journal Articles	# of Unique Keywords	Ratio of Unique Keywords	Average # of Keywords per Article
International Journals	*IPM*	4761	1045	3114	0.65	4.6
	JAL	1693	351	1091	0.64	4.8
	JASIST	736	278	345	0.47	2.6
	JD	3262	700	1353	0.41	4.7
	JIS	3491	665	2452	0.70	5.2
	Scientometrics	11420	2246	5237	0.46	5.1
Korean Domestic Journals	*JKBSLIS*	2495	499	1825	0.73	5.0
	JKSLIS	4119	909	2768	0.67	4.5
	JKSIM	2435	486	1839	0.76	5.0
	JKLISS	3814	802	2734	0.72	4.8

hand, the *JKSIM* (0.76) has the highest ratio of unique keywords. In general, Korean LIS journals have a higher ratio of unique keywords than the international journals. In general, it is reasonable to assume that the more duplicate keywords are found in a dataset, the greater the chance should be that the ratio of unique keywords is lower. However, the ratio could be affected by various factors such as normalization accuracy differences between Korean and English keywords, cultural and linguistic differences in choosing keywords, etc. Thus, although international journals contain higher ratio of unique keywords than Korean journals, more rigorous investigation is needed to examine this phenomenon.

2. Frequently Used Keywords in Each Journal

The frequency count of the keywords was performed on a per journal basis in order to obtain their relative ranks. Table 3 shows the result of ranking the most frequently used keywords of each journal. In Table 3, general keywords and specific keywords co-occur regardless of journal type. Some keywords such as "information retrieval" and "bibliometrics" are regarded as core research areas in the LIS community. On the other hand, keywords such as "internet" can be viewed as a subject area but lack an additional qualifying term (e.g, technology). Some keywords such as "evaluation" and "assessment" can be considered as general terms that lack an additional qualifying term. Thus only the most frequently used keywords in this table perfectly match the subject areas of LIS.

In Table 3, a variety of keywords that could be considered as a legitimate subject areas of LIS research can be observed. The result also demonstrates the fact that the journals share common keywords in the top frequency count list, implying that LIS journals often publish research papers in the same subject areas. In this table, duplicate instances, where the keywords that appear more than once across the journals, are shown in bold. As the list expands, because some higher ranked keywords is likely to appear in multiple journals, non-bold keywords is likely to change into bold letters. At the same time, some new lower ranked keywords is likely to be added to the list with non-bold letters, because the lower the rank the less they will be duplicated across journals.

<Tab. 3> Top 10 Most Frequently Used Author Supplied Keywords in International Journals

International Journals					
Rank	*IPM*	*JAL*	*JASIST*	*JD*	*JIS*	*Scientometrics*

Rank	*IPM*	*JAL*	*JASIST*	*JD*	*JIS*	*Scientometrics*
1	**information retrieval**	academic libraries	**bibliometrics**	**information retrieval**	**information retrieval**	**bibliometrics**
2	evaluation	**information literacy**	**citation analysis**	information science	**knowledge management**	**citation analysis**
3	query expansion	library instruction	**information retrieval**	libraries	ontology	h-index
4	natural language processing	assessment	**information seeking**	user studies	knowledge sharing	**scientometrics**
5	machine learning	open access	text mining	classification	metadata	citations
6	text categorization	academic librarians	**scientometrics**	internet	**information literacy**	bibliometric analysis
7	**citation analysis**	scholarly communication	natural language processing	information management	**citation analysis**	impact factor
8	question answering	social media	**knowledge management**	information	world wide web	collaboration
9	**information seeking**	university libraries	informetrics	information literacy	data mining	co-authorship
10	digital libraries	academic library	information use	information research	sentiment analysis	research evaluation

Korean Journals			
Rank	*JKBSLIS*	*JKSLIS*	*JKSIM*	*JKLISS*

Rank	*JKBSLIS*	*JKSLIS*	*JKSIM*	*JKLISS*
1	**public library**	**public library**	network analysis	**public library**
2	**library**	**library**	**library**	**Library**
3	**school library**	**school library**	ontology	information literacy
4	**academic library**	reading education	**public library**	**public libraries**
5	reading	**academic library**	metadata	**school library**
6	bibliotherapy	teacher librarian	citation analysis	library management
7	user satisfaction	**university library**	co-word analysis	**Librarian**
8	**public libraries**	bibliotherapy	academic libraries	**academic library**
9	**librarian**	kdc	research trends	**university library**
10	type of bad behavior	topic map	user study	Metadata

Focusing on the similarities and differences, two of the most frequently used keywords − "information retrieval" and "bibliometrics" − appear to be common throughout the international journals. The most frequently used keyword in *JASIST* and *Scientometrics* is "bibliometrics", while "information retrieval" is the most frequently used keyword in *IPM*, JD, and JIS. The most frequently used keyword list in *JAL*, which is "academic libraries", is an exception in this regard. As the name *JAL* (*Journal of Academic Library*) suggests, the most frequently used keyword is

"academic libraries". Clearly, the focus of the journal is reflected in the keywords, especially in the most frequently used words list. Also, the most frequently used keywords are most appropriate to be used as core LIS research areas. Similarly, in Scientometrics, bibliometrics is the highest used author keyword due to the focal point of the journal.

To characterize keywords of the international journals even further, most keywords shown in this table are roughly in line with the major theme of the journal, as described in Table 1. For example, *JAL* deals with topics related to academic library. To this end, most of keywords found in the *JAL* in this column is closely associated with "library". Similarly, the keywords found in *Scientometrics* appear to be narrowly focused to an extent. Thus, the keyword list under *Scientometrics* is closely related to the subject areas of bibliometrics.

In the lower portion of Table 3, striking differences and similarities among Korean journals in terms of keyword characteristics can be noticed. The most frequently used keyword is "public library" in Korean journals. This keywords is ranked at the top in three of the four Korean journals - *JKBSLIS, JKSLIS,* and *JKLISS*. The keyword "public library" is still respectably ranked #4. As suggested by Seo et al. (2015), this result implies that research on public libraries is most common in Korea. In general, three Korean journals − *JKBSLIS, JKSLIS, JKLISS* − show a remarkable similarity in terms of common keyword usage. *JKSIM* is more distinctive among the Korean journals, in the sense that fewer instances of duplicate keywords across journals are present in the list. Furthermore, *JKSIM* contains more keywords that specifically refer to methodology (e.g., network analysis). However, the methods can be applied to the library domain, and, to this end, the keyword "public library" is still ranked high (#4) in the list.

However, when compared to the international journals, this keyword does not appear in the top 10 most frequently used keyword list. In the international journal side, some keywords such as "information retrieval" are most frequently used throughout the journals as more than half of LIS international journals contain the keyword. In contrast, none of the top 10 keywords in the Korean journals contains both "information retrieval" and "bibliometrics". Oddly, even though *JKSIM* deals with areas related to information science, these keywords do not appear in the top 10 list.

3. Frequently Used Keywords in Two Datasets: International and Korean

The most frequently used keywords in international journals can be compared to Korean

journals in a number of ways. It can be argued that using a dataset that consists of simple aggregated keywords leads to bias toward a particular journal since one journal may publish many more articles than another. As an alternative approach, a ranking of frequency of keywords was first obtained independently for an international journal dataset and for an Korean journal dataset. Then, the keywords that co-exist across all journal types (i.e., international and Korean) were obtained, disregarding the keywords that do not appear in another journal. Because some keywords become eliminated in this process, the rank of frequently used keywords in each journal dataset was re-calculated based on the sum of each journal rank of frequently used keywords.

For example, in Table 4, "public library" is ranked #1 in *JKBSLIS*, *JKSLIS*, and *JKLISS,* and #4 in *JKSIM.* Since the sum of all these rankings (1 + 1 + 4 + 1) is equal to 7, this summed value was used to re-assess the total rank. After combining the ranking across the journal type, re-ranking was applied based on the combined rank of each journal. This ensured that the keywords were used throughout the journals and not in just one journal. The highest ranking means that the total rank would be the lowest, and having the lowest rankings would mean the total rank would be the highest, and vice versa. The result of using this scheme is shown in Table

⟨Tab. 4⟩ Frequently Used Keywords in Two Opposing Datasets

Total Rank (Combined Journal Rank)	Korean Journals	Total Rank (Combined Journal Rank)	International Journals
1	public library (7)	1	bibliometrics (173)
2	library (8)	2	evaluation (209)
3	academic library (38)	3	citation analysis (292)
4	school library (39)	4	scholarly communication (406)
5	public libraries (42)	5	knowledge management (565)
6	university library (53)	6	twitter (841)
7	information literacy (64)	7	content analysis (948)
8	academic libraries (72)	8	classification (1241)
9	metadata (82)	9	collaboration (1380)
9	library management (82)	9	spain (2346)
9	user study (82)	9	performance (2814)
12	user satisfaction (90)	12	social sciences (2883)
13	information service (114)	13	innovation (2413)
14	digital library (114)	14	h-index (3804)
15	librarian (118)	15	research methods (4225)

Note: The number in the parenthesis represents the combined ranking from individual journals.

4. This table depicts the frequently used keywords that co-exist across Korean journals, and the frequently used keywords that independently co-exist in international journals. Considering Korea and international as two independent domains, two datasets were created on the basis of keywords that appear across journals. As shown, most keywords appear to represent some common subject areas of LIS. The keywords that do not sufficiently indicate the subject areas of study lacks specificity as they are too general in terms of being able to be represented as a subject area. For instance, the keyword "evaluation" needs another qualifying word in order to increase its specificity.

Apart from these types of keywords, the most frequently used keyword that appear across international journals is "bibliometrics". Some keywords related to bibliometrics can be grouped together due to being conceptually close to one other. For example, two keywords - "citation analysis" and "h-index" - are closely associated with bibliometrics. In contrast, the most frequently used keyword across Korean journals is "public library".

Compared with Table 3, in Table 4, the most frequently used keyword across the international journals is "information retrieval", but the keyword "information retrieval" does not surface in Table 4 since an entirely different method was used to create the two datasets. Although the subject areas of information retrieval predominantly appeared in some international journals, this might be an indication that articles published in the subject area of information retrieval might not be best suited in some LIS international journals. For instance, the subject areas of information retrieval appear to have little relevance to the subject areas of scientometrics, and journal articles on information retrieval is less likely to be published in the journal *Scientometrics*. In contrast, the result indicates that the subject area of bibliometrics appears to be well-suited for all types of international LIS journals.

As discussed previously, the overarching theme in the Korean journals is the library. Although a variety of themes can be noticed in the international journals, a notable theme in the international journals is the bibliometrics. The keywords "bibliometrics" is ranked #1 in terms of frequency count and other closely related keyword such as "citation analysis" and "h-index" rank within the top 10. In general, the frequently used keywords represent similar and divergent subject areas between the two datasets.

Admittedly, this simple aggregation of keywords is prone to sampling bias due to the use of a dissimilar number of keywords pertaining to the journals. For example, in Table 2, we have

seen that *Scientometrics* publish a much greater number of journals than *JASIST*, and the keywords show a greater dissimilarity due to this difference in the number of published articles. Consequently, analyzing the dataset, which consists of only six international journal articles, may not be all-encompassing and should be considered in conjunction with other procedures described in this article.

4. Distinctive Keywords in Each Journal

Identifying distinctive keywords will aid in discovering the subject areas that are particularly unique to the international and Korean journals. The advantage of identifying distinctive keywords is that we may find keywords that closely match the unique subject areas of a particular domain. Moreover, as will be discussed later, unique traits of keyword usage can be easily discovered. A number of ways can be devised to determine the distinctiveness of keywords. The process of assessing and selecting distinctiveness depends on the definition of "distinctiveness" (Ridell 2014).

For the purpose of this study, the distinctive keywords are referred to as the keywords that appear in one domain, but not in another. For example, we can assess the distinctiveness of keywords per journal or per journal type − Korean or international. In Table 5, the distinctive keywords of each journals are shown. For the six international journals, the distinctive keywords are detected by discriminating the keywords that appear in the respective international journals but not in others. Similarly, the distinctive keywords of each Korean journals are detected by discriminating the keywords that appear in the respective journal but not in others.

Table 5 shows the top 10 most frequently used distinctive keywords in each journal. To obtain this result, a ranking was performed on the frequency count of keywords appearing in one journal but not in others. As a whole, most keywords reflect specific areas of LIS research. Some keywords in Table 5 can be directly associated with the highly ranked keywords that were shown in Table 3. For instance, "language model" in *IPM* can be considered as a secondary subject area that is closely associated with "information retrieval" which is the most frequently used keyword. These two keywords do not share a common vocabulary but are conceptually close to one another.

With the main focus of each journal in mind, by examining this keyword list, we can gain additional insights to the type of published works pertaining to each journal. For instance, keywords such as "church library" indicates a unique subject area, but are in line with the major

theme of the Korean journal dataset, which is the "library". The most frequently used keywords in *JKSIM* is "network analysis". Interestingly, even though this is a short list, some frequently used unique keywords pertaining to *JKSIM* can be associated with the keyword "network analysis" (e.g. web mining, Rocchio algorithm, and intellectual analysis).

〈Tab. 5〉 Top 10 Frequently Used Distinctive Keywords

Rank	*IPM*	*JAL*	*JASIST*	*JD*	*JIS*	*Scientometrics*
1	distributed information retrieval	library instruction	computer mediated communications	information research	digital divide	nanotechnology
2	data fusion	academic librarians	organization of information	public libraries	online information retrieval	research assessment
3	context	university libraries	technology impact	information searches	social informatics	indicators
4	language modeling	academic library	knowledge modeling	worldwide web	project management	university
5	Svm	information literacy instruction	information science history	classification schemes	polarity classification	patent analysis
6	human information behavior	library services	information resources management	literacy	user needs	Italy
7	probabilistic model	international students	automatic extracting	documents	use statistics	publication analysis
8	language model	faculty	tree structures	individual behaviour	topic maps	triple helix
9	term weighting	college students	reviewing	modelling	tags	patents
10	similarity search	mentoring	personal information systems	communication technologies	subject metadata	research output

Rank	*JKBSLIS*	*JKSLIS*	*JKSIM*	*JKLISS*
1	Blog	tagging	libqual+	information policy
2	user needs	one city one book reading campaign	web mining	knowledge information
3	reading culture program	property	service evaluation	economic feasibility
4	library volunteers	Impact	rocchio algorithm	contracting out
5	digitization	five laws of library science	retrieval interface	collaborative repository library
6	consortium	citation order	research evaluation	university students
7	CIDOC CRM	user service model	profiling	the library
8	church library	undergraduate	intellectual structure analysis	teaching competency
9	university archives	transmission	world wide web	teacher librarian education system
10	type of bad behavior	topic map	user study	Subcontracting

Also, it is worth noting that scholarly communication is considered as one of the notable subject areas in LIS. This keyword is not present in any of the Korean journals. It is possible that no work in this area has been published in the Korean journal. Although scholarly communication can be considered an independent LIS research area, the frequently used distinctive keywords, in general, appear to reflect the overall theme of the journal as the keywords are often in line with the most frequently used keywords that were identified earlier.

5. Distinctive Keywords in Two Datasets: International and Korean

A comparison can be also made between the set of distinctive keywords drawn from the two opposing datasets - one dataset created from the Korean journal keywords and another dataset created from the international journal keywords. The distinctive keywords of a dataset refer to keyword that appears in one dataset but not in another. As shown in Table 6, some distinctive keywords surfaced when this procedure was performed. As pointed out earlier, the principal theme in Korean journals is the library. Some distinctive keywords in this list also carry a similar theme (e.g., school library). In international journals, bibliometrics is a still major theme as previously pointed out. Although the actual keyword "bibliometrics" was not present in this list, keywords closely representing bibliometrics surfaced in the list (e.g., scientometrics and webometrics). Thus some keywords in each dataset carry similar themes as the top-ranked distinctive keywords of the dataset. When the two datasets are compared to each other, the issues relating to the school library appears to be common in the Korean LIS research but not in the international LIS research. In essence, using this method, traces of the unique themes concerning Korea can be found.

In Table 6, the most frequently used keyword in the international journal dataset is "scientometrics". This keyword does not appear in the Korean journal dataset. As shown, "bibliometrics" is a most frequently used keyword in international journals. Similarly, "scientometrics" carries the theme of bibliometrics, and this keyword is detected as a distinctive keyword in the international journal dataset. However, no instances of "scientometrics" are found in the Korean dataset.

The fact that one dataset is domestic and the other dataset is international provides a plausible explanation for this phenomenon. That is, there could be a subtle cultural bias toward the

selection of academic vocabularies. This is evident in the use of the keyword "teacher librarian", as shown in Table 6. The term "teacher librarian" is widely accepted in Korea, but less in elsewhere. The term "school librarian" or "media specialists" is preferred in most regions of the world. Such anecdotal evidence suggests that the use of keywords tends to be bound by a specific domain (e.g., a specific nation or region) due to regional differences in the usage of keywords.

Also, there could be differences in keyword usage due to a unique national concern that the keyword represents. For example, KDC, which refers to less likely to appear in the international journal dataset. In sum, such traits of keyword usages imply that critical examination of frequently used keywords in one community versus another should be considered in characterizing the works of LIS research. It appears that the distinctive keyword identification procedure described in this study seems to be valuable in discovering domain-specific keywords.

〈Tab. 6〉 Top 10 Distinctive Keywords in Korean Journals and International Journals

Rank Based on Frequency Count	Distinctive Keywords inKoreanJournals	Rank Based on Frequency Count	Distinctive Keywords inInternationalJournals
1	school library (106)	1	scientometrics (112)
2	bibliotherapy (42)	3	peer review (52)
3	teacher librarian (37)	4	webometrics (50)
4	reading education (36)	5	search engines (36)
5	information service (35)	6	ranking (35)
6	library policy (29)	6	publications (35)
7	KDC (27)	6	nanotechnology (35)
8	library service (22)	8	sentiment analysis (33)
9	reading program (21)	9	patent (33)
10	reading guidance (18)	10	research assessment (31)

Note: The number in the parenthesis represents the frequency count from each dataset.

V. Summary and Conclusion

This study so far examined similarities and differences between Korean journals and international journals in terms of keyword uses. The results of this study can be summarized using the following key points: a) substantial overlapping subject areas emerged in the three Korean journals: *JKBSLIS, JKSLIS, JKLISS*. b) Although one of the most frequently used keywords in

JKSIM is "library", it is evident from the keyword analysis that the journal also touches upon many core subject areas of information science. c) The most frequently used keyword in more than half of the international journals was "information retrieval", but no instances of "information retrieval" was found in the Korean journal dataset. d) A frequently used distinctive keyword such as "scholarly communications" was detected in the international journal dataset, but no instances were found in the Korean journal dataset. e) It was evident that some keywords are bounded by nation or by a region (e.g., KDC) because of its scope of applicability. f) Based on the ratio of unique keywords, more polarized keywords appears to be present in the international journals than in the Korean journals.

With respect to Korean journals and international journals, this article has demonstrated that the most frequently used keywords － including frequently used distinctive keywords － appear to accurately reflect the various subject areas of LIS research. Although some subject areas, such as academic libraries, are covered in both international and Korean journals, this study demonstrates that numerous distinctive subject areas are not covered in both international and Korean journals. Korean LIS research is dominated by library related research. In contrast, LIS research in the international arena tends to be more diverse, although there is an international LIS journal such as *JAL* that specifically focuses on the academic library related subject matters. The keyword analysis also demonstrates that, journal articles on information retrieval are much more common in international side than the journal articles on information retrieval in Korean side.

After conducting this study, the randomly selected international journals appears to be somewhat inadequate in representing the influential LIS international journals. The selection of journals is prone to bias sampling, and the result of this study to be generalized to the entire population, which is all international journals, is limited. For instance, if an LIS journal called *Library and Information Science Research* was selected instead of *Scientometrics,* then the result could be different in terms of the frequently used keywords, the distinctive frequently used keywords, etc. As the name suggests, *Library and Information Science Research* is more general than *Scientometrics,* and it is still ranked within the top 30 *Scimago SJR*. Thus, this limitation of the study should be considered when interpreting the results of this study. To this end, additional analytical research on keywords are desirable, particularly with larger, diverse collections of datasets.

Despite this limitation, the noteworthy implication of this study is that a critical analysis of

keywords seems to be effective in understanding the domain of interest. An extensive list of distinctive keywords used in a specific LIS community could be identified using the relatively simple procedure described earlier. Such research efforts should facilitate an understanding of the LIS field as a whole, and, as a result, the researchers in LIS communities should be able to identify new research opportunities by reviewing common and less common subject areas of the journals.

References

Barthel, Roland and Roman Seidl. 2017. "Interdisciplinary Collaboration Between Natural and Social Sciences−Status and Trends Exemplified in Groundwater Research." *PloS One,* 12(1): e0170754.

Bawden, David and Lyn Robinson. 2015. *Introduction to Information Science.* Facet Publishing.

Cunningham, Scott W. and Jan Kwakkel. 2011. "Innovation Forecasting: A Case Study of the Management of Engineering and Technology Literature." *Technological Forecasting and Social Change,* 78(2): 346-357.

Dougherty, Dale and Arnold Robbins. 1997. 2nd Ed., *Sed & Awk: UNIX Power Tools.* O'Reilly Media, Inc.

IPM Home Page. 2017. <https://www.journals.elsevier.com/information-processing-and-management> [cited 2017. 2-19].

JAL Home Page. 2017. <https://www.journals.elsevier.com/the-journal-of-academic-librarianship> [cited 2017. 2-19].

JASIST Home Page. 2017. <http://onlinelibrary.wiley.com/journal/10.1002/%28ISSN%2915322890/homepage/ProductInformation.html> [cited 2017. 2-19].

JD Home Page. 2017. Emerald Publishing. Journal of Documentation Information. <http://www.emeraldgrouppublishing.com/products/journals/journals.htm?id=jd> [cited 2017. 2-19].

JIS Home Page. 2017. <http://journals.sagepub.com/home/jis> [cited 2017. 2-19].

JKBSLIS Home Page. 2017. <https://kbiblia.jams.or.kr/co/main/jmMain.kci> [cited 2017. 2-19].

JKLISS Home Page. 2017. <https://liss.jams.or.kr/co/main/jmMain.kci> [cited 2017. 2-19].

JKSIM Home Page. 2017. <http://www.koreascience.or.kr/journal/AboutJournal.jsp?kojic=JBGRBQ> [cited 2017. 2-19].

JKSLIS Home Page. 2017. <http://www.koreascience.or.kr/journal/AboutJournal.jsp?kojic=MHJBB6> [cited 2017. 2-19].

Lee, Seongsin. 2016. "A Study on Research Trends in Public Library Research in Korea Using Keyword Networks." *Libri,* 66(4): 263-274.

Li, Wentian. 1992. "Random Texts Exhibit Zipf's-Law-Like Word Frequency Distribution." *IEEE Transactions on Information Theory,* 38(6): 1842-1845.

Milojević, Staša, Cassidy R. Sugimoto, Erjia Yan and Ying Ding. 2011. "The Cognitive Structure of Library And Information Science: Analysis of Article Title Words." *Journal of the American Society for Information Science and Technology,* 62(10): 1933-1953.

Névéol, Aurélie, Rezarta I. Dogan and Zhiyong Lu. 2010. "Author Keywords in Biomedical Journal Articles." In *AMIA Symposium Proceedings:* 537-541.

Newman, Mark E. 2005. "Power Laws, Pareto Distributions and Zipf's Law." *Contemporary Physics,* 46(5): 323-351.

Riddell, Allen B. 2014. "Feature Selection: Finding Distinctive Words" In TAToM: *Text Analysis with Topic Models for the Humanities and Social Sciences.* Göttingen: DARIAH-DE. <https://de.dariah.eu/tatom/feature_selection.html>. [cited 2017. 2-19].

Scientometrics Home Page. 2017. <http://link.springer.com/journal/11192> [cited 2017. 2-19].

Seo, Eun-Gyoung, Won-Kyung Lee, Eun-Kyung Park and Ock-Seong Lee. 2015. "Informetric Analysis of Research Trends in the Journal of Korean Biblio Society for Library and Information Science." *Journal of the Korean BIBLIA Society for Library and Information Science,* 26(3): 315-343.

Zhang, L. Ling and Hyun Jin Hong. 2014. "Examining the Intellectual Structure of Reading Studies with Co-Word Analysis Based on the Importance of Journals and Sequence of Keywords." *Journal of the Korean Biblia Society for Library and Information Science,* 25(1): 295-318.