

LIS Research in Japan

Shoichi TANIGUCHI, Keita TSUJI, Fuyuki YOSHIKANE
(Univ. of Tsukuba)

Table of contents

1. Introduction
2. Analysis of research papers published in domestic journals
3. Analysis of research papers published in international journals
4. Analysis of grant-winning research projects
5. Conclusion

1. Introduction

The purpose of this paper is to show the current situation of Library and Information Science (LIS) research activities in Japan, more precisely, to show an overview of research paper production and research projects in the LIS field in Japan. First, the background on research activities will be described briefly, which we should know before discussing the research activities. Then, the results of an analysis will be shown of research papers published in domestic journals and also of papers in international journals. After that, the results of an analysis on research projects that win grants will be touched. Finally, the main findings will be reviewed.

There should be several issues influencing research activities themselves or analysis on such activities in Japan. The important issues of that kind would be the followings:

(1) Weakness in LIS education system. Japanese LIS education can be categorized into two types: LIS professional education and “shisho” training, which means qualified, certificated librarian. Fewer than 10 universities have LIS program, which is at undergraduate, master, and/or doctoral level. They have sufficient faculty staff to pursue research. In contrast, more than 150 colleges and universities provide shisho training course. In almost all cases that course is operated with one or two full-time faculty staff members and several part-time lecturers. They are not in sufficient situation to pursue research.

(2) Weakness in libraries' or other related organization. Only a small number of large libraries (national or university ones) have research section and staff of their own. This leads to the fact that only university teachers have time and money to pursue research.

(3) Multi-/inter-disciplinarity of LIS field. Multi-/inter-disciplinarity is an intrinsic feature of the LIS field itself and implies vague boundary between other fields. This issue is related to difficulties in conducting a survey or analysis of all LIS research activities. The scope of the LIS field varies depending on researcher or database, and hence a consistent survey or analysis is sometimes difficult. This is likely to be a common issue with other countries.

2. Analysis of Research Papers Published in Domestic Journals

2.1 Analysis with BIBLIS Database

BIBLIS is the most comprehensive bibliographic database covering domestic journals in LIS field, and it was created and maintained by the Japan Society of Library and Information Science. However, the maintenance of the database was suspended in 2007, because of an economic issue. We thus adopted BIBLIS database and limited to the scope of 1991 to 2006, which is 16 years long. We also limited the scope to 24 current major journals and their preceding ones, and to the document type “ research paper ” by referring to the document type code assigned to individual documents. The database contains a variety of documents and journals. But it is enough to see the major journals and their research papers for showing an overview of research activities in LIS field.

The 24 major journals that we selected for the analysis are as follows:

● Library and Information Science	93 papers
● 日本図書館情報学会誌 (J. of JSLIS)	187
● TP&D フォーラムシリーズ (TP&D Forum Series)	50
● 現代の図書館 (Libraries Today)	381
● 情報の科学と技術 (J. of Info. Sci. and Tech. Association)	898
● 情報管理 (J. of Info. Proc. & Management)	547
● 図書館界 (The Library World)	366
● 大学図書館研究 (J. of College and Univ. Libraries)	298

The first and second ones are the most important. The first one is *Library and Information Science* published by the Mita Society for Library and Information Science. During the 16 years, 93 research papers in total have been published in that journal. The second one is *Journal of the Japan Society of Library and Information Science* published by that society, which is also the creator of BIBLIS database. The journal has published 187 research papers in total during that period. These two journals are typical peer-reviewed ones.

On the other hand, most of journals listed here have published lots of papers. For example, about 900 papers have been published in the *J. of Information Science and Technology Association* and more than 500 in the case of *J. of Information Processing and Management*. Papers of those journals are peer-reviewed in some cases and not peer-reviewed in most cases. That is the reason for lots of papers being published.

2.2 Basic Statistics on Papers and Authors

The basic statistics from this analysis are the followings:

- In total 4,591 papers have been published in 24 journals during the 16 years, including both peer-reviewed and not peer-reviewed papers. Almost all papers were written in Japanese language with some exceptions.
- The average number of papers per year was 287; nearly 300 papers have been produced every year. We think this number indicates sufficiently active states of LIS research in Japan.
- For authors, 6,244 persons and corporate bodies were counted in total, which indicates 1.36 authors per paper. This means about 1.4 times of the papers equals the number of authors producing them. The authors include researchers, librarians and other practitioners. Researchers include university teachers and graduate students, in particular, doctoral course students

Figure 1 displays how many papers have been published each year. The lowest line designated with ' ' displays the number of papers produced in that year. Nearly 300 papers have been produced constantly, except years like 1991, 1998, and 2004.

In the figure, the top line with '· - ·' shows the total number of authors who published those papers. The middle line with '····' shows the total number of different, unique authors publishing them. We can observe a consistent trend: these three lines are

parallel to each other during the period.

In Figure 2, the top line designated with '—' is the average number of authors per paper. Regarding the average number of papers per author, we have two ways to calculate it. The first one is to divide the total number of authors by the number of different authors. The second one is to divide the number of papers by the number of different authors. The middle line with '· · · · ·' in the figure indicates how many times each author has occurred in the whole set of papers, which corresponds to the first measurement. The total average during this period is about 1.17 times. On the other hand, the lowest line with '· · - · ·' indicates how many papers each author has produced for that year, which is calculated in the second way. The total average is 0.86 papers. The yearly-variation of these numbers is relatively small.

Another research group conducted a similar, but more elaborate survey in 2006 (Mine et al., 2006). They obtained the result which was 0.81 papers per author. That preceding research result roughly matches that being obtained in this analysis.

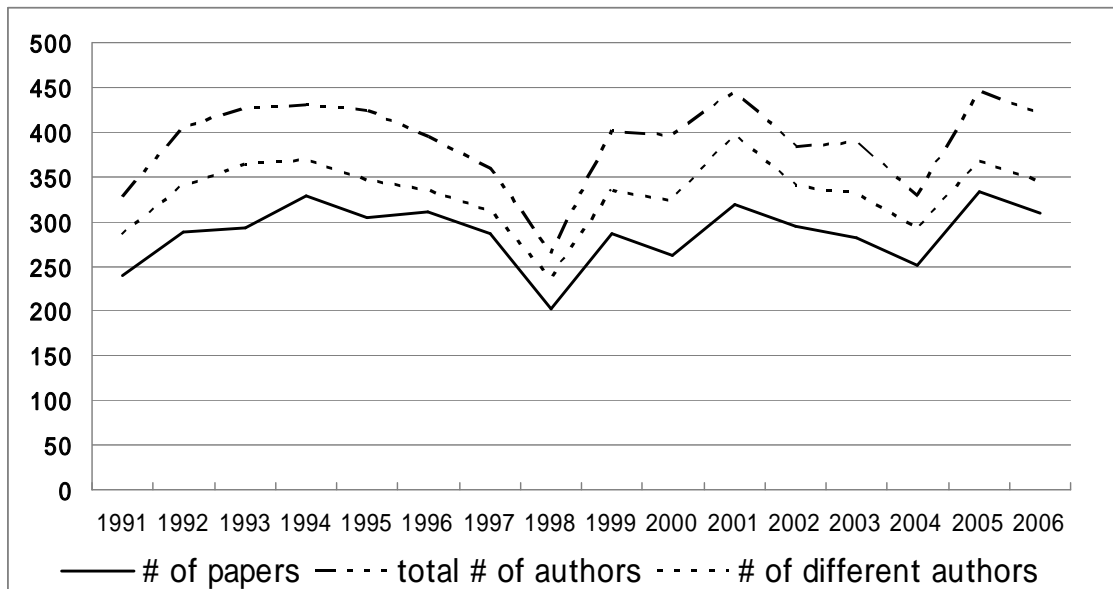


Figure 1. Number of papers and authors in 24 major journals.

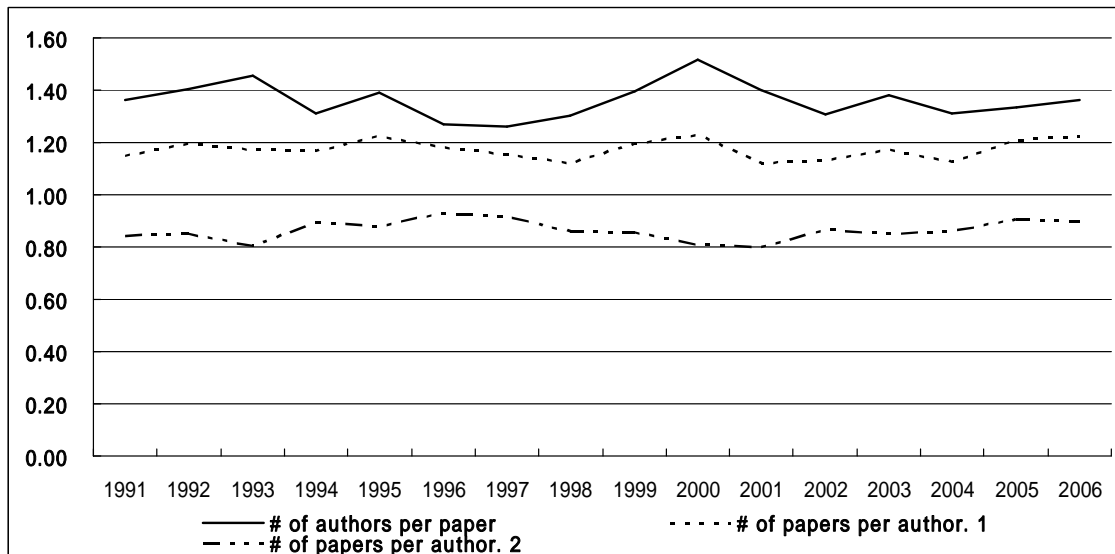


Figure 2. Number of authors per paper and of papers per author.

2.3 Topics by Classification

Next, we try to sum up those research papers based on BIBLIS classification code. BIBLIS has its own classification scheme, and individual papers in the database have a classification code assigned by the editors of the database.

The classification scheme is a hierarchical one. The top-level classes are as follows:

- A: general
- B: libraries & information centers (including policies, administration, architecture)
- C: media
- D: information organization
- E: information technology
- F: user studies, user behaviors
- G: information retrieval

Figure 3 shows the results of the analysis. 10% of the whole are categorized into class A: general. 34 % of the papers are categorized into class B: libraries & information centers, which were the largest group. 8 % is under the class C: media; 17% is under the class D: information organization. The class E: information technology, the class F: user studies and user behavior, and the class G: information retrieval are assigned 14%, 7%, and

10%, respectively.

The ratio of each class in each year is illustrated in Figure 4. The class B: libraries & information centers takes the largest percentage every year during the period.

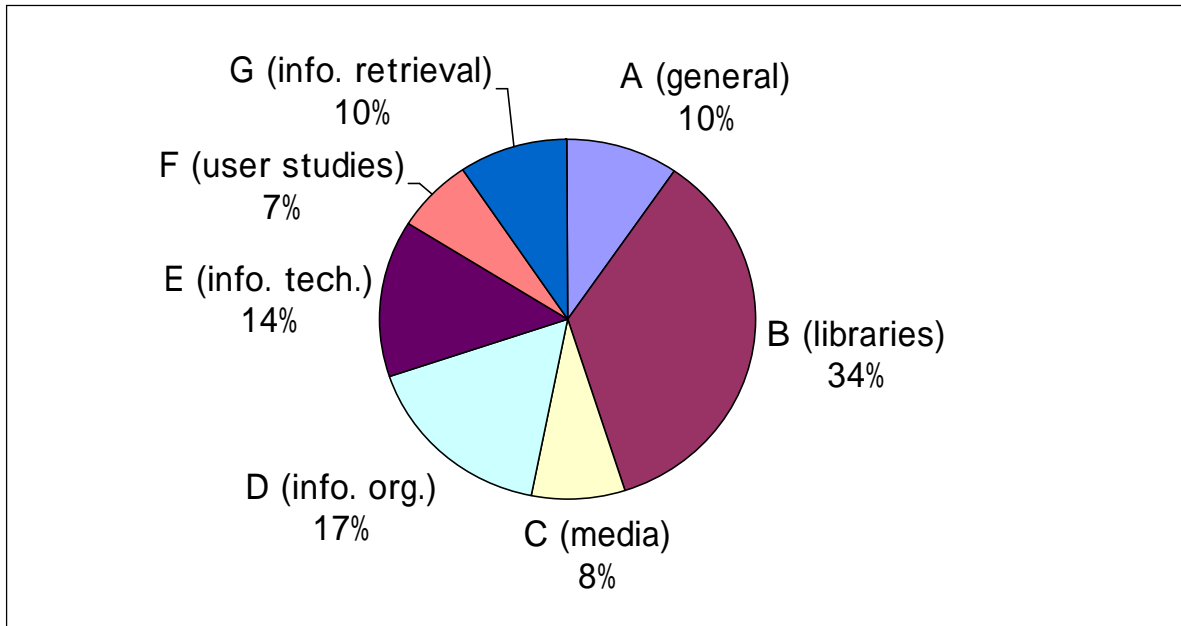


Figure 3. Percentage of classes for the whole period.

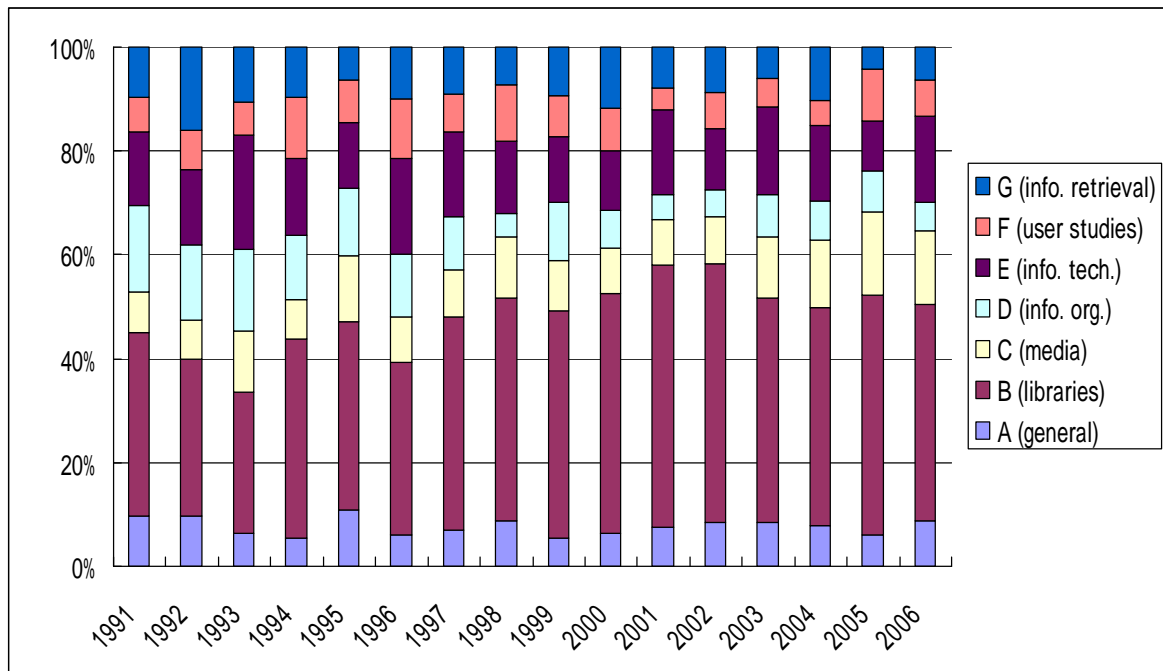


Figure 4. Percentage of classes in each year.

3. Analysis of Research Papers Published in International Journals

3.1 Analysis with Web of Science

In this section, we focus on research papers in international journals by Japanese authors. We took the database “Web of Science” for this analysis. It is the most comprehensive citation index for every field, as you know. Of course, major international journals in LIS field have been covered.

We used the database and limited the scope to the subject field “Information Science & Library Science” set out in the database. In addition, we limited it to the document type “article, proceedings paper, or review” and also to the publication year from 1982 to 2009. Of course, only English journals were adopted.

After that, within the set of papers satisfying the above conditions, we identified papers written by authors in Japan, by identifying authors whose address or affiliation was Japan. This means that Japanese people with foreign country’s address or affiliation were not included here.

As a result, 307 papers were identified. Within these papers, more than 20% have international coauthors, that is, coauthors in foreign countries. 20 papers (6.5%) have coauthors in USA, and 10 papers (3.3%) have coauthors in China. Canada (7 papers; 2.3%), England (5 papers; 1.6%), South Korea (5 papers), and others, are counties of other international coauthors.

The followings are the top eight journals publishing papers by Japanese authors. The top four journals take nearly 50% of the whole.

● INFO. PROCESSING & MANAGEMENT	59	19.2%
● SCIENTOMETRICS	47	15.3%
● J. OF THE AMERICAN SOC. FOR INFO. SCI. AND TECHNOLOGY (JASIST) including J. OF THE AMERICAN SOC. FOR INFO. SCI. (JASIS)	23	7.5%
● TELECOMMUNICATIONS POLICY	23	7.5%
● INTL. FORUM ON INFO. AND DOCUMENTATION	19	6.2%
● INTL. J. OF GEOGRAPHICAL INFO. SCI.	18	5.9%
● J. OF INFO. SCI.	8	2.6%

● J. OF THE AMERICAN MEDICAL INFORMATICS ASSOCIATION	7	2.3%
--	---	------

The followings are journals citing those papers by Japanese authors:

● SCIENTOMETRICS	91	9.5%
● LECTURE NOTES IN COMPUTER SCI.	47	4.9%
● J. OF THE AMERICAN SOC. FOR INFO. SCI. AND TECHNOLOGY (JASIST) including J. OF THE AMERICAN SOC. FOR INFO. SCI. (JASIS)	45	4.7%
● INFO. PROCESSING & MANAGEMENT	33	3.5%
● INTL. J. OF GEOGRAPHICAL INF. SCI.	30	3.1%
● LECTURE NOTES IN ARTIFICIAL INTELLIGENCE	19	2.0%

Journals citing those papers by Japanese authors were diverse; a wide variety of journals were listed. The top four journals accounted for less than a quarter of the entire citations.

Meanwhile, as for citation counts analysis on 307 papers by Japanese authors, 903 times in total they have been cited by other research papers within the scope of “Web of Science.” On average each paper has been cited 2.94 times. The h-index, which is used recently and also based on citation counts, is an index to designate activity of producing important papers. The h-index value of 307 papers is 14, which means 14 papers have been cited 14 or more times.

3.2 Comparison among Three Countries

Let us take a brief look at the activities in China and South Korea. The numbers were counted by use of “Web of Science” database, and the scope of the analysis was the same as that of Japan.

We identified papers written by authors in China, by limiting to authors whose address or affiliation is China. But, in this case, authors having Hong Kong address or affiliation before 1997 were not included. As a result, 798 papers were identified as such. Within these papers, about half of the papers have international coauthors. 20% of them (161 papers) have US coauthors, and 4% (35 papers) have Belgium and another 4% have

Canadian coauthors. Singapore (26 papers), England (22 papers), and others, are countries of other international coauthors. This number probably shows the very active international co-authorship and collaboration.

On the other hand, on papers by Korean authors, we identified 398 papers in total whose authors' address or affiliation is South Korea. More than 30% of the papers have international coauthors. 20% of them (80 papers) have US coauthors, and 3% (12 papers) have Chinese coauthors.

In Figure 5, the line with ' ' shows the number of papers written by Japanese authors. In recent years, a gradual increase can be observed. However, when we compare papers by Japanese authors with those by Chinese or South Koreans, we should notice an important fact: China and South Korea have produced more papers in international journals, and their growth rates are obviously higher than that of Japan.

Figure 6 displays the total cited counts of those papers in each year. These counts include the citations for all preceding papers published so far. Thus the counts probably increase year-by-year up to a given point, although other factors also affect the counts. For example, papers published in 2009 can cite all preceding papers published before, whereas papers in 2000 can cite all papers published before that time.

If we compare citation counts of papers by Japanese authors with those by Chinese or South Koreans, we get the fact that the gap between them has been getting bigger in recent years.

Japan has produced 307 papers in total during the period, and each paper has got 2.94 average citations, as we mentioned earlier. In contrast, China produced more than double the papers in the same duration, and average citation count per paper is 4.73. South Korea has produced 398 papers, but average citation per paper is 7.41, which is more than two times higher than that for Japan. For h-index, Japan is 14, China is 34 - it is a great number - and South Korea is 24. Table 1 shows these numbers for the three countries.

Let us summarize this section as follows:

- The amount of published papers in major international journals has been

growing in Japan, in particular, in recent years. However, the growth rate is small. In contrast, China and South Korea have shown a very rapid increase in the numbers of published papers.

- According to the numbers of international co-authorship, China and South Korea have been active in collaborative research with foreign countries' researchers, especially, US researchers. Japan has not been active yet in such international collaboration.
- From the citation analysis, both average citation counts per paper and h-index are low in Japan. Contrary to this, the other two countries have higher value than that of Japan in both average citations per paper and h-index. In particular, h-index in China and average citations in South Korea is very high.

These lead to a conclusion that research activities in those two countries are very active and impactful.

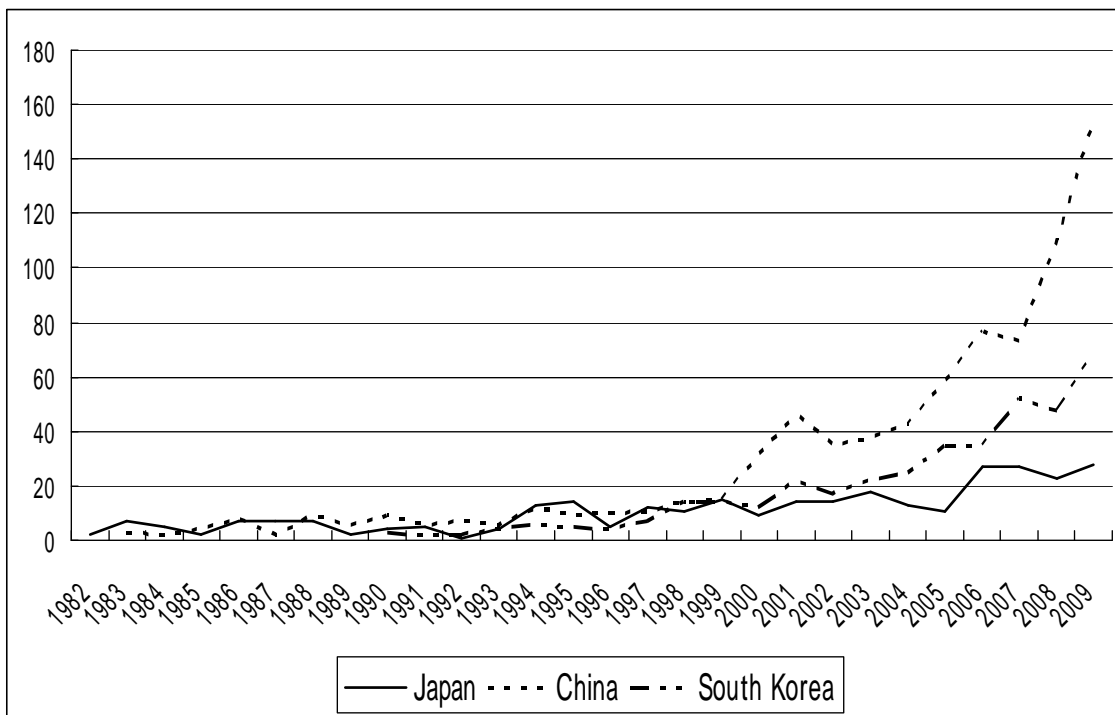


Figure 5. Number of papers published in international journals.

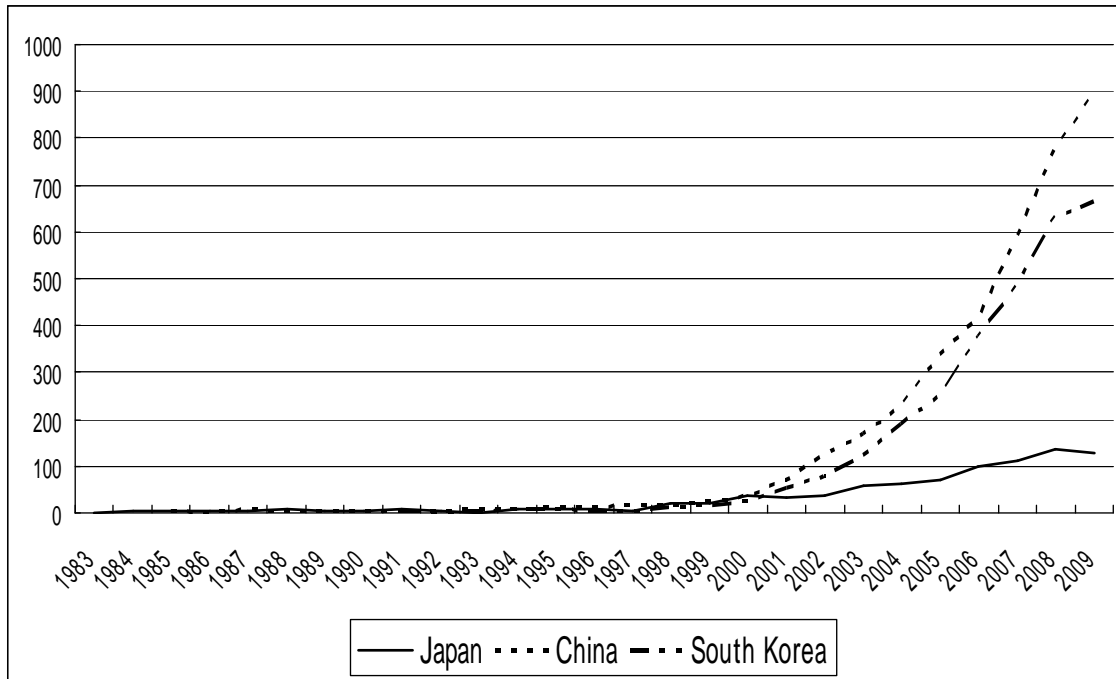


Figure 6. Citation counts for papers published in international journals.

Table 1. Comparison among three countries.

	# of papers	sum of the times cited	average citations per paper	h-index
Japan	307	903	2.94	14
China	798	3,775	4.73	34
South Korea	398	2,949	7.41	24

4. Analysis of Grant-winning Research Projects

4.1 Analysis with KAKEN Database

Now, we try to show some features of research projects getting grants in Japan. In Japan, we have several types or categories of research grants. Grants-in-Aid for Scientific Research offered by Ministry of Education, Science, Sports and Culture is the most popular for university researchers and hence for this analysis we used the database KAKEN which contains all project information about funded grants. Project

name, project years, researcher names, keywords, abstract, etc. are recorded in the database.

We got data on projects from the database which satisfies the following conditions:

- projects during 2003 to 2008 fiscal years.
- projects accepted in the research field category “LIS, Humanities informatics and Social sciences informatics.” We cannot divide the category into subcategories and thus cannot distinguish LIS-field projects from the other two. This is the limit of this analysis.
- projects having keywords assigned by project leaders. We removed three projects with no keywords.

Consequently, 264 projects were obtained. Some projects were large-sized, whereas most of them were small or middle-sized. Also, most of them are multiple-years projects and some are one-year projects. The average is 2.1 years per project. Some of the projects getting the grants in 2008 have continued in the succeeding years 2009 and 2010. If we sum up the number of projects in each year, we will get 559 projects in total.

For the projects until 2002, we did not use project data, because a different research field categorization had been adopted at that time and it was difficult to distinguish LIS- field projects from those in computer science, which were very numerous.

Table 2 shows the number of projects in each year. In 2003 only 49 projects got the grants, and in 2004, 77 projects were accepted. But recently more than 110 projects have got the grants. Such growth could be expected in future.

Table 2. Number of projects and keywords assigned.

Year	2003	2004	2005	2006	2007	2008
Projects	49	77	98	109	112	114
Keywords per project	13.8	12.9	12.8	11.1	9.2	6.9
Projects per keyword	1.11	1.11	1.14	1.14	1.16	1.17
Different keywords	611	900	1,096	1,063	886	671

4.2 Keywords Analysis

Individual project data contains keywords assigned by its project leader herself/himself, and those keywords represent the topic and features of the projects. In this analysis with KAKEN database, we focus on the above keywords assigned by project leaders.

More than 2,500 keywords in total were used in the projects. On average 9.67 keywords were assigned to each project. Each keyword was used only 1.31 times during the period. This implies a wide variety of keywords were used in order to characterize individual projects.

Table 2 displays such counts on keywords in each year. During the period 2003-2008, keywords per project have decreased year-by-year. In 2003, individual projects had 13.8 keywords on average. On the contrary, in 2008 each project had 6.9 keywords. We do not have any reason to explain this tendency yet. It might be the change of instructions for project leaders to assign keywords. But we have not been able to make sure about that point yet.

The following counts were obtained after English keywords were translated into Japanese whenever possible. In addition, the keywords “LIS,” “informatics,” and similar ones, which represent the discipline itself, were excluded from this analysis. Finally 1,462 different, unique keywords were identified.

The most frequent keyword was “database” (“データベース”), which has been used in 54 projects during the period. Other frequent used keywords are as follows - hereafter words in Japanese are not noted.

- metadata (34 projects), internet (30), public libraries (29), digital/electronic libraries (29), information systems (23), digital archives (19), GIS (geographical info. systems) (18), bibliometrics (18), archeology (17), libraries (16), ubiquitous (15)

Middle-range frequency keywords are listed below. Actually, keywords used in ten projects, nine projects, and eight projects are listed here:

- Used in ten projects (9 keywords): archive, simulation, portals, metadata schema, school libraries, info. service, info. society, university libraries,

character recognition.

- Used in nine projects (7 keywords): identities, USA, questionnaire surveys, citation analysis, bibliography, info. seeking behavior, info. systems on history.
- Used in eight projects (8 keywords): media, risk, co-authorship analysis, retrieval systems, researchers, e-government, computerization, social info. systems.

Keywords used in only one or two projects are enumerated below. Words with specific meaning or marginal words for LIS-field are displayed:

- Used in two projects (528 keywords): dspace, IC cards, ocr, rf-id tags, svm, web archives, web mining, xml-edi, xml messaging, innovation theory, image analysis, internet connectedness, interface, interaction design, ...
- Used in one project (173 keywords): sns, topic map, web2.0, xml retrieval systems, Arabic characters, wearable computers, ethnomethodology, governance, globalization, ...

Likewise, 454 keywords were used in three projects, 140 keywords in four projects, and 47 keywords in five projects. That is to say, lots of low-frequency keywords occurred.

The top line in Figure 7 indicates the number of projects getting grants in each year. The other lines in the figure represent the frequencies of most frequent keywords listed earlier. It follows from the figure that there is no keyword which has been used in proportion to the increase in project numbers.

Figure 8 shows the number of projects with most frequent keywords, which have been taken up in Figure 7. The keyword "database" occurred in 12 projects in 2003. However, in 2008 it occurred in 6 projects. We think that keyword has been widely accepted gradually in LIS field and became a general word which does not have a specific meaning to represent a research topic or feature. As the figure shows, those frequent keywords have been used constantly during the period.

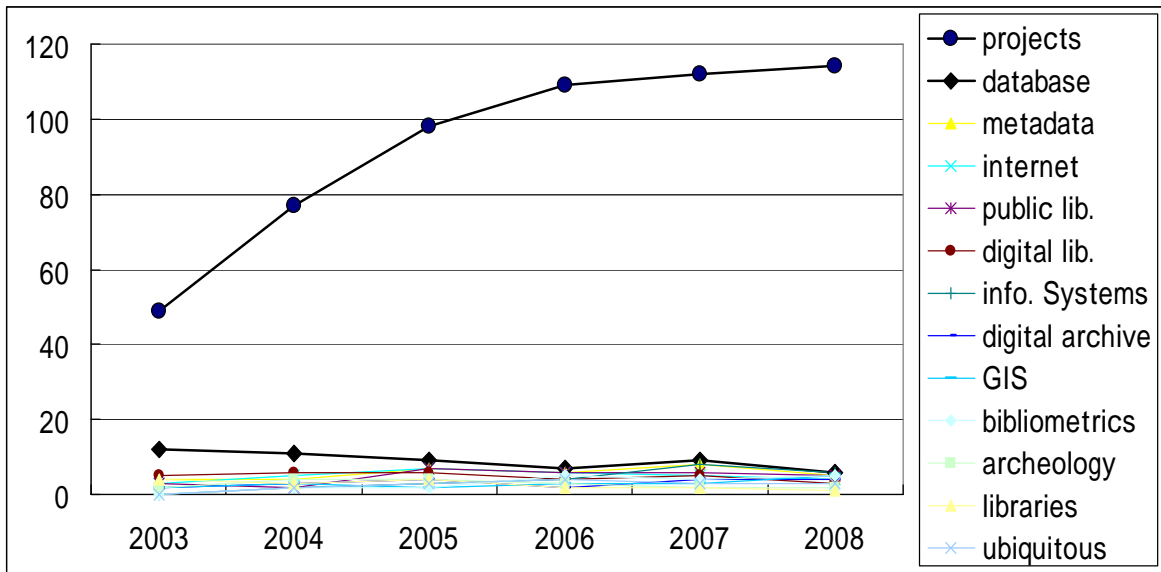


Figure 7. Number of projects and most frequent keywords.

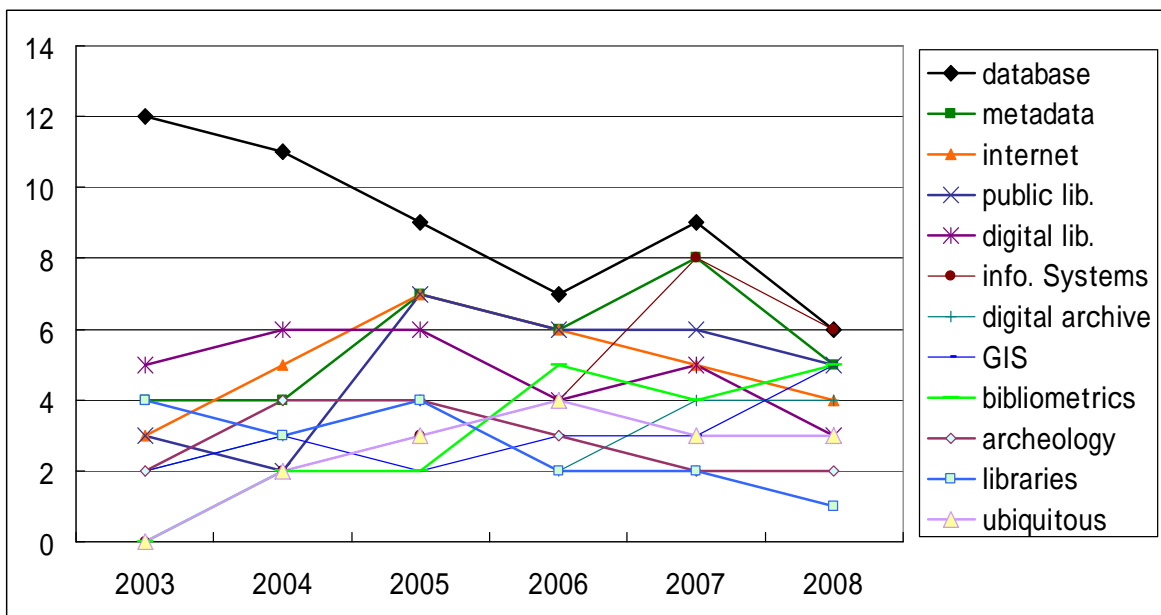


Figure 8. Number of projects with most frequent keywords.

4.3 Projects Distribution and Clustering

Additionally we tried an analysis on the research projects with MDS (Multi-dimensional scaling), a method for exploring similarities or dissimilarities in data. For this, MDS analysis was conducted with keywords representing the topics and features of the

projects. And hence keywords used in only one project were excluded. The aim of this analysis is to investigate the project distribution of different years and to make sure whether there is any explicit trend or not.

Figure 9 shows the result of MDS analysis for the projects of 2003, 2005, and 2008. The characters “3,” “5,” and “8” represent individual projects of 2003, 2005, and 2008, respectively. Distance between the characters indicates closeness or similarities between the projects corresponding to the characters. As the figure illustrates, these three characters are closely plotted and not able to discriminate each other. This implies that the projects of those three years are not discriminated each other with the keywords.

If we try to conduct the same analysis for the projects of different years' combination, like the projects of 2003 and 2008, or those of 2005 and 2008, we obtain the similar results: the projects of those two years are not well discriminated from each other. That is to say, no explicit difference was observed among the projects of different years. We think the period 2003-2008 was too short to identify a research trend.

Meanwhile, MDS analysis shows that the projects can be divided into several groups based on closeness or similarities. Hence, we also tried to apply clustering to the projects. Clustering is a method to set projects into clusters, each of which have closeness or similarity based on keywords. There are several ways of clustering and we adopted a method: hierarchical clustering with Euclidean distance and Ward's criterion.

Figure 10 is a result of the clustering for the projects of 2008. In this case, the projects were divided into 12 clusters. We tried to interpret the scope of the individual clusters built, but we realized it was tough.

If we venture to do that, the cluster 1 ('C1' in the figure) could be characterized with keywords “information service,” “academic libraries,” “public libraries,” and so on. The next cluster C2 could be characterized with “citation analysis,” “bibliometrics,” “library service,” etc. The largest cluster C3 has a wide variety of keywords. The core keywords would be “metadata,” “information technology,” “data mining,” “social media,” and so on. We are not sure why such a large cluster was built, and hence a closer investigation would be needed.

From the clustering of the projects within a year like 2003, 2005, or 2008, we realized that it was tough to identify the scope of individual clusters, as we mentioned earlier.

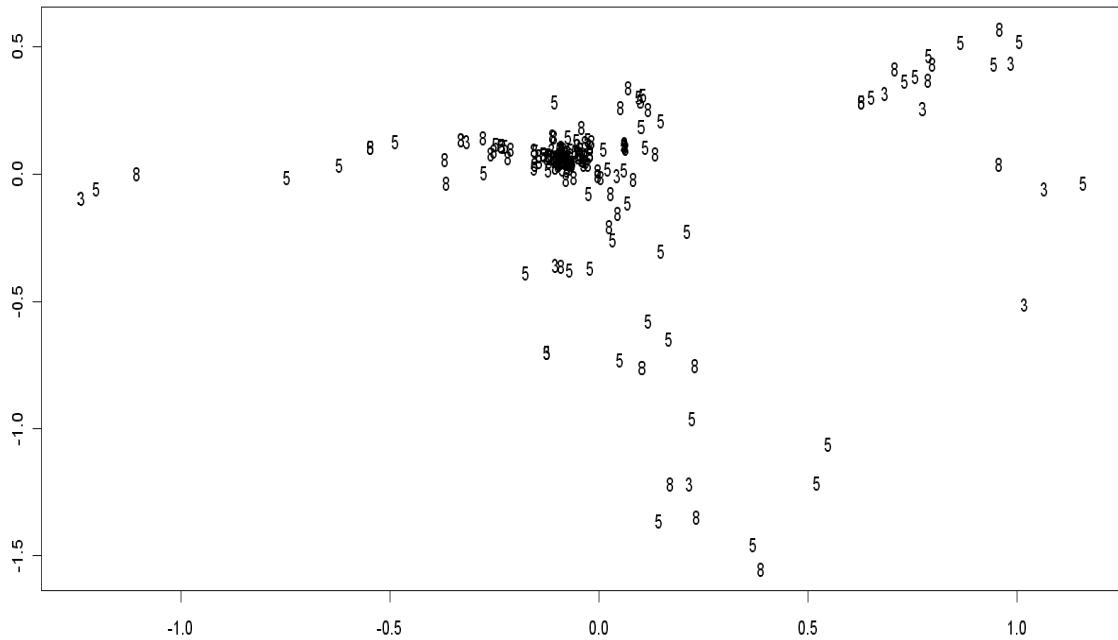


Figure 9. Project distribution of 2003, 2005, and 2008 with MDS.

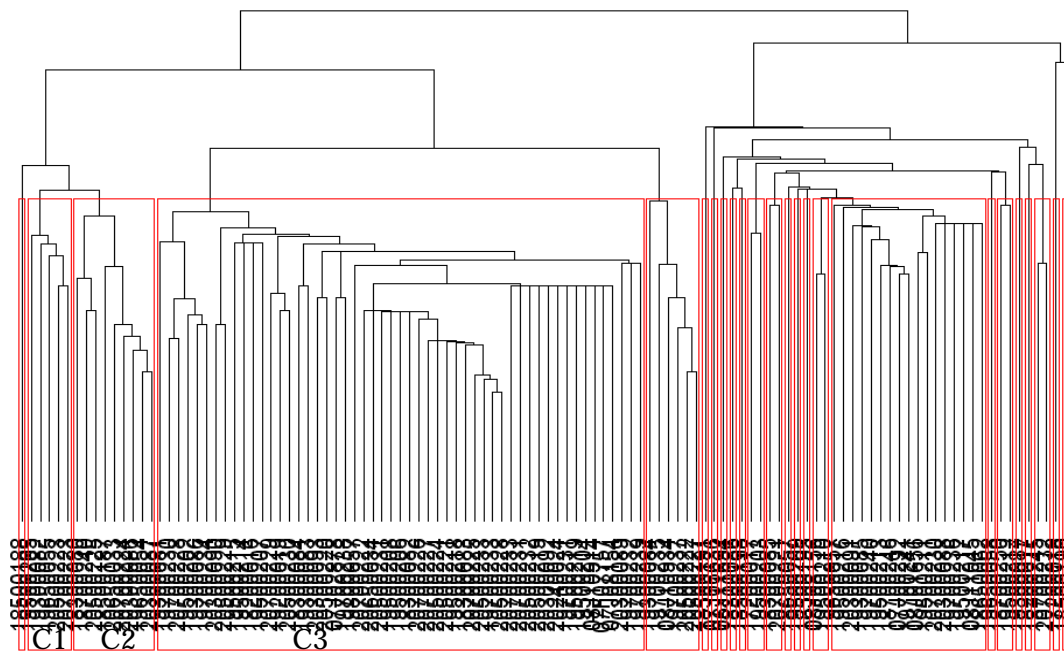


Figure 10. Projects clustering for 2008 (using 12 clusters).

5. Conclusion

For Japan, we observed the powerful activities in publishing research papers in domestic journals. In contrast, a limited number of research results have been published in international journals. We would like to say that LIS research in Japan is not active in the international community and not sufficient in international contribution.

What is the future choice that Japan should take? More globalization in research? Or, remain as a Galápagos Island? This is most likely the same issue that LIS education in Japan has been facing (Nemoto, 2009).

Acknowledgments

We wish to express our gratitude to Prof. Nobuyuki Midorikawa for give us this opportunity to present the results of our analysis. We would like also to thank the Japan Society of Library and Information Science for permission to use BIBLIS data in our analysis.

References

- Shinji Mine, Shuichi Ueda, Makiko Miwa. (2006). Library and information science educators in Japan: academic qualifications and research productivity. *Library and Information Science*. No. 55, p. 71-82. (in Japanese) URL:
<http://mslis.jp/pdf/LIS055071.pdf>
- Makiko Miwa, Shuichi Ueda, Akira Nemoto, Mitsuhiro Oda, Haruki Nagata, Teruyo Horikawa. (2006). Final Results of the LIPER Project in Japan. 72th IFLA General Conference and Council, Education and Training with Library Theory and Research , 2006-08-20/24, Seoul, Korea. URL:
<http://archive.ifla.org/IV/ifla72/papers/107-Miwa-en.pdf>
- Akira Nemoto. (2009). Galápagos or an isolated model of LIS educational development?: a consideration of Japanese LIS education in the international setting. Symposium on Future Perspectives in Globalization of Library and Information Professionals: Asia-Pacific Conference on Library & Information Education and Practice(A-LIEP

2009). 12 p. URL: <http://a-liep.kc.tsukuba.ac.jp/proceedings/Symposium/005.pdf>

Research Committee, Japan Society of Library and Information Science. (1998).
Toshokan-jouhou-gaku Kenkyu to Sono Shien Taisei (LIS research activities and their
supporting system). Japan Society of Library and Information Science, 86 p. (in
Japanese) URL: http://wwwsoc.nii.ac.jp/jslis/publications/sien_taisei_1.html