Employing Informetric Analysis to Identify Dominant Research Areas in the Top Ranking U.S. LIS Schools

계량정보학적 분석을 통한 특정 대학원의 핵심 연구분야 파악: 미국 상위 10개 문헌정보학 대학원을 대상으로

Hae-Young Kim*
Ji-Hye Lee*
Young-Mee Chung**

ABSTRACT

Authoritative as well as objective information on ranking or dominant research areas of academic departments/schools in a certain discipline is essential for the graduate school applicants. In this study, we performed an informetric analysis to identify dominant research areas in the top 10 U.S. LIS schools. We used two different datasets of research productivity and research interests of the LIS faculty. The correspondence analysis method was employed to graphically display the association between research areas and the LIS schools. We found that the research productivity data collected from SSCI database generated a very informative map presenting which research areas were dominant in which LIS schools. We also found that for the two most productive subject areas in LIS over the past 10-year period, the proportion of research articles in information retrieval decreased to a great extent in the recent 5-year period, whereas that of information seeking behavior showed an almost same degree of increase.

초 록

이 연구에서는 대학원 지원자들이 특정 연구분야가 강한 대학원을 선택하는 데 도움을 주기 위해 어느 대학원이 어떤 연구분야에서 특히 강한가를 파악할 수 있는 계량정보학적 분석 방법을 제안하였다. 분석을 위해 미국의 상위 10개 문헌정보학 대학원을 선정하고, 각 대학원 교수진의 최근 10년간의 연구논문과 홈페이지에 등록되어 있는 관심분야를 수집하였다. 연구생산성 데이터가 관심분야 데이터에 비해 더 신뢰할 수 있는 분석 결과를 보였으며, 빈도 데이터 분석과 대응일치 분석 결과 각 대학이 어느 주제분야에서 특히 강한가를 파악할 수 있었을 뿐만 아니라 대학원 간의 인접성을 측정할 수 있었다.

Keywords: informetric analysis, correspondence analysis, research productivity, research areas, LIS schools
계량정보학적 분석, 대응일치분석, 연구생산성, 연구분야, 문헌정보학대학원

---

* The Graduate School, Yonsei University(siru@hanmail.net; cooljh99@naver.com)
** Professor, Dept. of Library and Information Science, Yonsei University (ymchung@yonsei.ac.kr)(Corresponding author)

■ Received : 21 May 2008 ■ Revised : 28 May 2008 ■ Accepted : 10 June 2008
■ Journal of the Korean Society for Information Management, 25(2): 143-155, 2008. [DOI:10.3743/KOSIM. 2008.25.2.143]

1. Introduction

Recently, the well-known bibliometric analysis methods such as citation analysis, co-citation analysis, and co-word analysis have been widely applied in analyzing various aspects of scholarly communication. Those studies include mapping of intellectual structure in scientific disciplines as well as science in general, detecting research fronts and collaboration patterns in certain research community and comparing the patterns between disciplines as well as countries, and evaluation of researchers, research groups, and institutions using various bibliometric indicators. In more recent years, webometric analyses have been performed to explore the topics such as subject structure of academic Web spaces, comparison of citation and link analyses to e-journals, link structure analysis of national as well as international academic spaces, and comparison of webometric indicators.

Data unit of analysis has also expanded from individual authors and documents to cover journals, departments, universities or other research institutions, and nations. Especially, the evaluation and ranking of research productivity of researchers, departments, and universities have attracted the attention of policy makers, administrators, research funding agencies, and graduate school applicants. For the evaluation of research productivity or research quality, previous studies have used such measures as scholarly publication counts, citation counts, impact factors, and h-index. For example, Meho and Spurgin(2005) reviewed 9 such studies in ranking LIS(Library and Information Science) schools in the United States. Among them, four studies used both publication and citation counts, one study citation counts only, and four studies publication counts only. Most of these studies used SSCI(Social Science Citation Index) database either alone or with additional databases such as LLIS(Library Literature & Information Science) and LISA(Library and Information Science Abstracts), and two of them used LLIS database only. Adkins and Budd(2006) performed another study to measure the scholarly productivity of U.S. LIS faculty using publication and citation data collected from the SSCI database. Meho and Spurgin (2005) examined the coverage of 9 LIS related databases with the publication lists of 68 faculty members of the top 10 ranked LIS schools and found out that no single database provided comprehensive coverage of the LIS literature. For journal articles the highest percentage of items covered by a single database was around 50% in comparison with about 84% coverage of all the databases.

The h-index, most recently proposed, has been used to evaluate scientific research in Spain (Imperial and Uez-Navarro 2007), and to evaluate chemistry research groups(van Raan 2006) and information scientists(Cronin and Meho 2006).

Authoritative and objective information on ranking or dominant research areas of graduate schools in a certain academic discipline is essential for the graduate school applicants who plan to continue their study in a very specialized field because they tend to select such a department or school that

has many prominent faculty members in a given research field. The purpose of this study is to suggest an informetric analysis method which could effectively identify research areas of strengths and weaknesses in a given graduate school by means of research productivity and specialties of the faculty. To this end, we analyzed two datasets representing faculty productivity and faculty specialties in the top 10 LIS schools selected from the ranking list of the US News Best Graduate Schools 2008 to discover which schools are strong in which subject areas. The approach taken in this study could be used for the graduate schools in other disciplines.

2. Methodology

2.1 Data Collection

In this study, we collected publication data of the top 10 LIS school faculty from the SSCI database, which has been used by most of the previous studies on LIS faculty productivity. We limited document type to journal articles and conference papers since SSCI covers only 1.8% of book chapters and no authored books(Meho and Spurgin, 2005). The top 10 LIS schools are University of Illinois at Urbana Champaign, University of North Carolina at Chapel Hill, Syracuse University, University of Washington, University of Michigan, Rutgers University, Indiana University, University of Pittsburgh, University of Texas at Austin, and Florida State University in rank order. Among the top 10 LIS schools, only 4 schools include the term "library"in their school names, e.g., School of Library and Information Science, and the remaining schools have names such as School of Information Science/Studies and School/College of Information. A few schools offer specific programs including Biological Informatics, Information Management, and Telecommunications and Networking.

To retrieve the relevant data, we searched address(AD=) and source(SO=) fields using "AND" Boolean operator, creating the dataset of 580 research articles published by researchers affiliated with the 10 LIS schools. Downloaded data included full bibliographic records for the two consecutive periods, 1998-2002 and 2003-2007. Source journal list includes 48 non-review LIS journals indexed by SSCI and 10 additional journals publishing LIS related articles. These 10 journals, classified under the categories other than Information Science & Library Science in JCR(Journal Citation Reports) database, were selected from the journals citing LIS journal articles more than 100 times. The additional journal list includes Decision Support Systems, European Journal of Information Systems, International Journal of Human Computer Studies, Journal of Computer Information Systems, Methods of Information in Medicine, Industrial Management Data Systems, Journal of Computer Mediated Communication, Libraries the Cultural Record, Computers in Human Behavior, and Information Retrieval.

To collect the data on faculty specialties, we

visited each LIS school's Web site in January 2008 and printed out the list of each faculty member's research or teaching interests as well as the recent courses offered by them. As a result, we obtained the specialty data on 261 full-time faculty members of the top 10 LIS schools. The specialty data collected were also used in identifying subject areas covered in the library and information science discipline.

2.2 Methods of Data Analysis

We employed the content analysis method to classify the collected data, i.e., publications and specialties, according to the subject areas reflecting the research or teaching interests of the LIS faculty. The subject area list was based on the classification scheme proposed by Meho and Spurgin(2005), but expanded to include 23 categories in total representing the research areas highly cited in the collected specialty data. The subject areas identified in this study include human-computer interaction, information-seeking behavior, information retrieval, information systems, mining & natural language processing, digital libraries, informetrics, collaboration, organization of information, information management, archives & preservation, management, reference & information services, public libraries, school library & children's literature, information policy, telecommunication & information technology, social informatics, special informatics, information literacy, computer-mediated communication, computer-assisted education & learning, and LIS general(see Table 1). Here, "information management" includes data management, warehousing, and knowledge management, whereas "management" includes library management and general management techniques. "Social informatics" is distinguished from "special informatics"in that social informatics deals with the social, political, and economic aspects of informatics, while special informatics includes informatics in special subject fields such as medical informatics in medicine. "LIS general"includes research methodology, LIS education, librarianship in general, philosophy, history, and others.

Regarding the classification of publications, we assigned a subject area to a research article on the basis of the title, keywords, and abstract in each document record. Although some papers could be classified into more than one subject area, we selected only one dominant area. On the other hand, research interests of each faculty member were classified into the matching subject areas resulting in multiple subject areas for each member.

To present the analysis result in a graphic form, we used the correspondence analysis method. Correspondence analysis is a method of analyzing cross-tabular data in the form of numerical frequencies and displaying the result in a simple graphical space which maps their association in two or more dimensions(Greenacre, 1993). The correspondence analysis method has been used in analyzing bibliometric data to compare the publication output patterns of different countries(Dore et al. 1996; Dore and Ojasoo 2001) or to reveal international collabo-

ration patterns(Anuradha and Urs 2007). In this study, we employed the correspondence analysis method to map the association between the individual LIS schools and the subject areas in library and information science by analyzing the faculty publications and research interests. We used the correspondence analysis program provided in the SPSS version 12.0 for mapping.

3. Results and Discussion

We had two datasets to be analyzed, i.e., the publication dataset and the research interest dataset. The former included 580 journal articles published by the faculty of each LIS school for the period of 1998-2007, consisting of two subsets covering two consecutive 5-year time-spans. The latter included the total of 709 research interests of 261 LIS faculty members that were used in identifying major subject areas in the LIS discipline.

3.1 Analysis of the Publication Data

<Table 1> shows the numbers as well as percentages of research articles in each subject area for the two 5-year periods plus the whole 10-year period. From <Table 1>, we can notice the two most productive subject areas for the 10-year period are information retrieval(14.7%) and information seeking behavior(10.9%). However, the two subject areas demonstrate the opposite trend over the past 10 years, that is, the significant increase of publications in information seeking behavior(from 6.7% to 12.5%) and the decrease in information retrieval(from 18.8% to 13.0%), thus resulting in almost the same proportions for the recent 5-year period. We also discovered noticeable changes in two other subject areas; social informatics area shows an increase from 1.2% to 3.9%, whereas mining & natural language processing area shows a decrease from 4.2% to 0.7%. This phenomenon may reflect the recent trend that the research on "hard"subject areas such as information retrieval and mining based on natural language processing has been diminishing in the LIS research community and the papers in those areas have been published more frequently in conference proceedings than in scholarly journals.

The data table in <Appendix 1> presents the frequency and two percentage values, i.e., external strength and internal strength, of the classified research articles of individual LIS schools. Regarding the school productivity, top three schools are North Carolina University(14.1%), University of Washington(13.3%), and Syracuse University(12.8%), and the lowest two schools are University of Michigan (5.3%) and University of Texas(5.7%). Regarding the area productivity, the top five productive areas are information retrieval(14.7%), information seeking behavior(10.9%), reference & information service(7.9%), informetrics(7.1%), and organization of information(6.6%) in frequency order.

To understand the table, consider the "information seeking behavior" cell of University of Illinois. The faculty of the University of Illinois has pub-

〈Table 1〉 Frequency Distribution of Research Articles in 23 Subject Areas

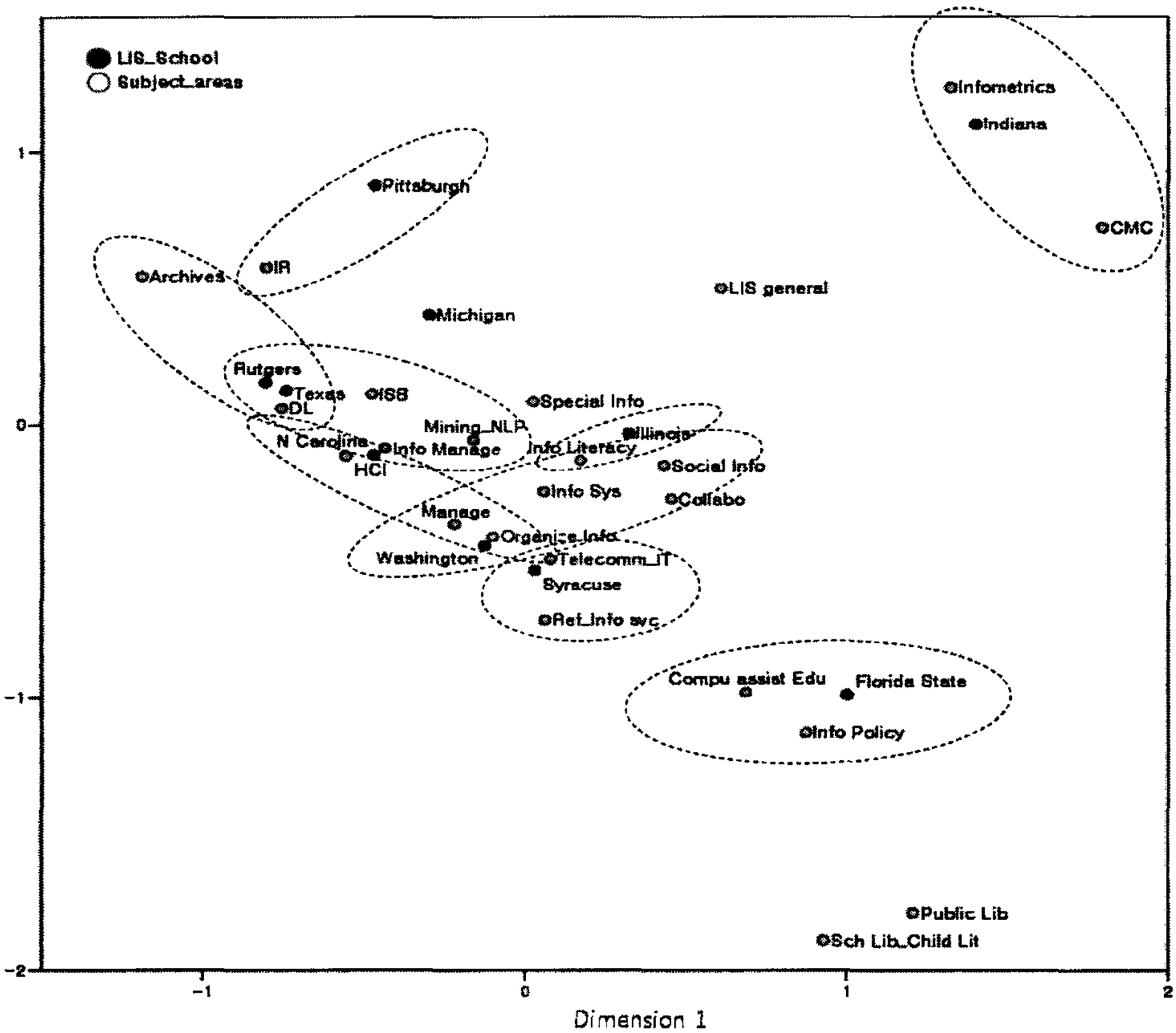
	Subject areas	first 5 years (1998 - 2002)		second 5 years (2003 - 2007)		Total (1998 - 2007)	
1	Human-computer interaction	10	6.1%	23	5.5%	33	5.7%
2	Information-seeking behavior	11	6.7%	52	12.5%	63	10.9%
3	Information retrieval	31	18.8%	54	13.0%	85	14.7%
4	Information systems	8	4.8%	16	3.9%	24	4.1%
5	Mining & Natural language processing	7	4.2%	3	0.7%	10	1.7%
6	Digital libraries	6	3.6%	13	3.1%	19	3.3%
7	Informetrics	13	7.9%	28	6.7%	41	7.1%
8	Collaboration	3	1.8%	13	3.1%	16	2.8%
9	Organization of information	13	7.9%	25	6.0%	38	6.6%
10	Information management	7	4.2%	11	2.7%	18	3.1%
11	Archives & preservation	0	0.0%	7	1.7%	7	1.2%
12	Management	7	4.2%	16	3.9%	23	4.0%
13	Reference & information service	11	6.7%	35	8.4%	46	7.9%
14	Public libraries	1	0.6%	8	1.9%	9	1.6%
15	School library & children's literature	0	0.0%	2	0.5%	2	0.3%
16	Information policy	3	1.8%	13	3.1%	16	2.8%
17	Telecommunication & IT	9	5.5%	16	3.9%	25	4.3%
18	Social informatics	2	1.2%	16	3.9%	18	3.1%
19	Special informatics	4	2.4%	10	2.4%	14	2.4%
20	Information literacy	4	2.4%	13	3.1%	17	2.9%
21	Computer-mediated communication	3	1.8%	11	2.7%	14	2.4%
22	Computer-assisted education & learning	2	1.2%	9	2.2%	11	1.9%
23	LIS general	10	6.1%	21	5.1%	31	5.3%
	Total	165	100.0%	415	100.0%	580	100.0%

lished 7 SSCI-indexed research articles, corresponding to 11.1% of the total research articles classified in this subject area. Again, the 7 articles in information seeking behavior occupy 16.7% of the total 42 articles published by the University of Illinois faculty. The percentage data mean that the University of Illinois is relatively strong in the information literacy area(23.5% of external strength) among the 10 LIS schools, while within the university, relatively dominant area is information seeking behavior(16.7% of internal strength) among the 23 subject areas. We call the first strength as external strength and the second as internal strength. In the table, the external values greater than 20% or the highest values are underlined for each subject area.

Regarding the subject areas with high external strength, we found the universities presumed strong

in certain subject areas with the external strength values greater than 30%. University of Indiana is strong in informetrics(43.9%) and computer-mediated communication(42.9%), University of Rutgers strong in Information management(38.9%) and relatively strong in mining & natural language processing(30.0%), University of Texas strong in Archives(42.9%), Florida State University very strong in public libraries(66.7%), and computer-assisted education(45.5%), University of North Carolina strong in special informatics(43%), and Syracuse University relatively strong in mining & natural language processing(30.0%), information policy(31.3%), and telecommunication & IT(32.0%).

<Figure 1> displays a biplot correspondence map of the subject areas and LIS schools based on the productivity data. The inertia, an eigenvalue for each dimension, is 0.223 for the first dimension and 0.142 for the second dimension. These values reflect the relative importance of the dimensions. The proximity of the dominant subject nodes and the school nodes generally corresponds to the origi-



〈Figure 1〉 Correspondence Map Based on the Research Productivity Data

nal data in <Appendix 1> with only a few exceptions. That is, subject nodes with high external strength for a given LIS school are positioned near the school node. For example, Florida State University node is located adjacent to information policy and computer-assisted education nodes, Indiana University node is adjacent to informetrics node, and so on. The correspondence map also demonstrates the proximity of the LIS schools with regard to the subject areas of strengths and weaknesses based on the research productivity. We can see that Indiana University, Florida State University, and University of Pittsburgh nodes are far apart from the rest of the LIS schools indicating that the dominant research areas of these schools are quite distinguishable from other schools.

3.2 Analysis of the Research Interests Data

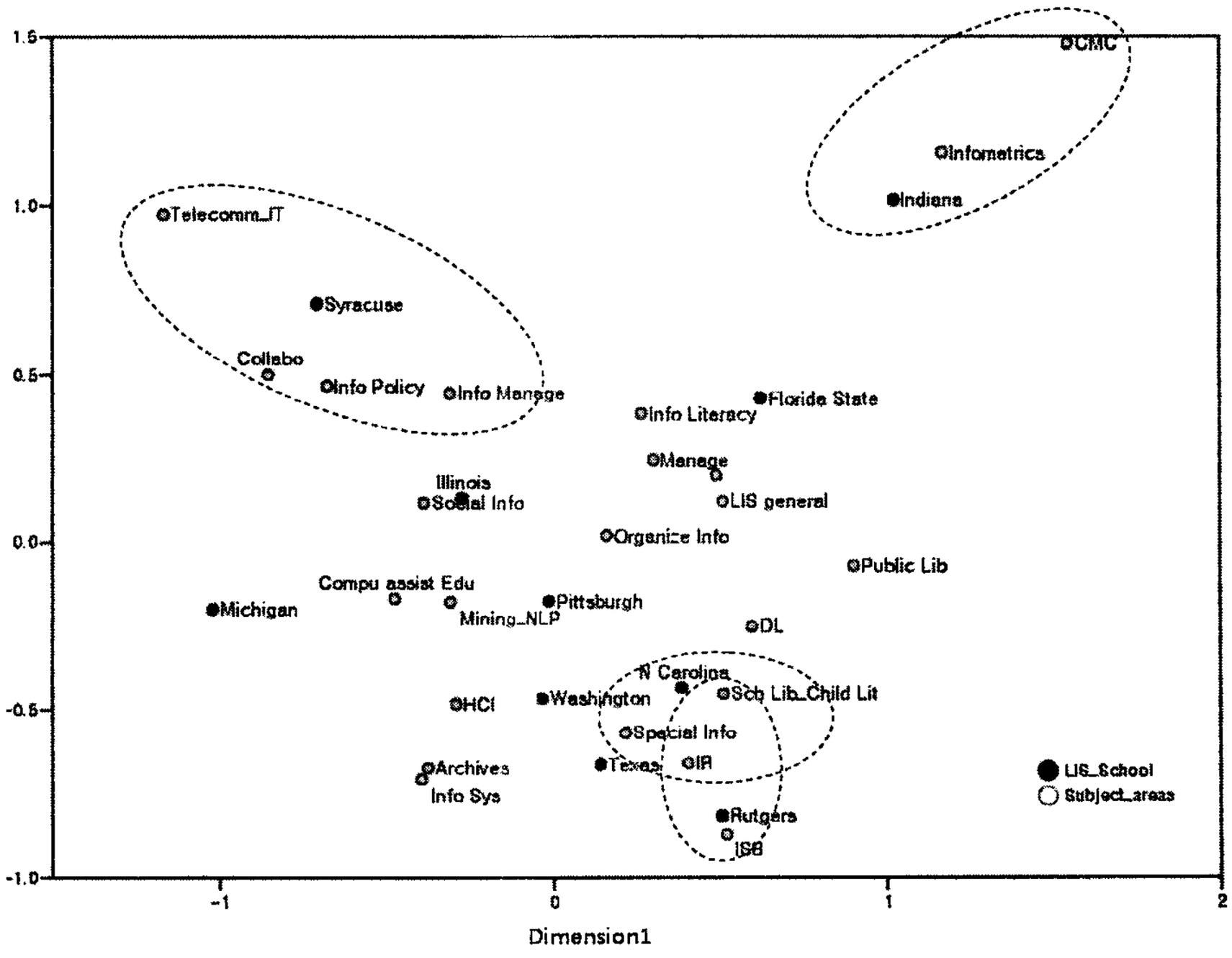
The data table in <Appendix 2> presents the frequency and two percentage values, i.e., external strength and internal strength, of the classified research interests of the faculty of individual LIS schools. The average number of research interests is 2.7, and the most frequently matched subject areas are social informatics(10.1%), LIS general(8.8%), information systems(7.7%), and HCI (7.2%). It is interesting to see that the two datasets, one created by research productivity and the other by research interests, show quite different frequency distributions. Furthermore, the subject areas of high external strength for individual LIS schools hardly correspond to those generated by research productivity. The best matching schools are Indiana University and University of Texas. The worst matching schools are University of Illinois, University of Pittsburgh, and Rutgers University. The rest of the schools show a partial matching. Florida State University presents a very conflicting data that is worthwhile to mention. In the research productivity analysis we found public libraries, computer-assisted education, and information policy as the most productive areas. However, it was found that the frequency of research interests was very low in information policy and computer-assisted education.

<Figure 2> shows the biplot correspondence map displaying the association between the subject areas and the LIS schools based on the research interests data. The inertia is 0.111 for the first dimension and 0.093 for the second dimension, both of which are much smaller than those of the map based on the productivity data. The correspondence map in <Figure 2> does not reflect well the subject areas of high external strength for each school as shown in the tabular data of <Appendix 2>. We can identify only four proximity clusters corresponding to the tabular data, which was presumably caused by the relatively small values of the high external strength, mostly less than 30% unlike those for research productivity as shown in <Appendix 1>. The most remarkable cases are: (1) Syracuse University node is surrounded by the four subject area nodes with external strengths greater than 20%, namely, telecommunication &

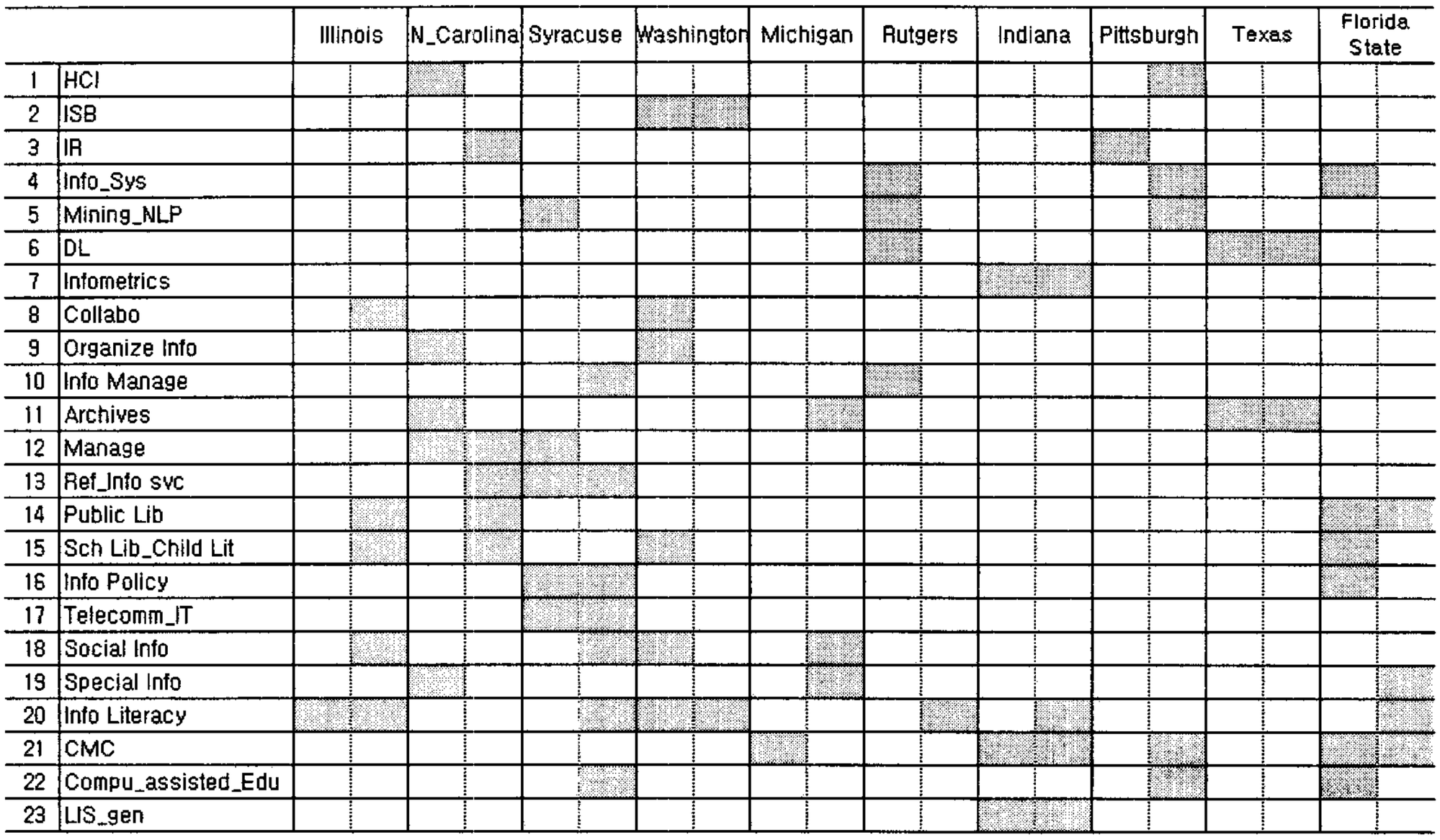
IT, information policy, collaboration, and information management, and (2) Indiana University node is isolated from other schools, very similar to the map based on the research productivity data, and is also positioned near the informetrics and computer-mediated communication nodes.

3.3 Discussion

<Figure 3> compares the prominent research areas identified by research productivity with those identified by research interests for each LIS school. The first column denotes the result from the productivity data and the second column denotes one from the specialty data. Darkened boxes indicate the subject areas with the external strength values greater than 20% or the largest value in the given subject area. The research areas of high strength in both research productivity and research interests are information literacy at University of Illinois, management at University of North Carolina, reference & information service, information policy, and telecommunication & IT at Syracuse University, information seeking behavior at University of Washington, informetrics and computer-mediated education, and LIS general at Indiana University, digital libraries and archives at University of Texas, and public library at Florida State University.



〈Figure 2〉 Correspondence Map Based on the Research Interests Data



〈Figure 3〉 Dominant Research Areas of the LIS Schools Identified from Research Productivity(first column) and Research Interests(second column)

4. Conclusion

In this study, we intended to identify the dominant research areas of the top 10 LIS schools in the United States using two different types of informetric data, i.e., faculty research productivity and faculty specialties. We employed the measure of external strength representing each school's relative strength in a certain subject area and the correspondence analysis method to map the association between the subject areas and the LIS schools.

We found that the dominant subject areas of each LIS schools identified by research productivity data were quite different from those identified by research interest data. We also found the research productivity data were more reliable in discovering the research areas of strengths and weaknesses in the LIS schools.

The limitation of the study is that the research productivity data were collected from the SSCI database only, thus resulting in an incomprehensive dataset of research productivity. However, we demonstrated the usefulness of informetric analysis in investigating the research areas of strengths and weaknesses in the LIS schools and more specifically, the usefulness of the correspondence analysis method in revealing the association between the research areas and the LIS schools as well as in displaying the proximity of the LIS schools with regard to the dominant research areas.

References

Adloms., D. and Budd, J. 2006. "Scholarly productivity of U.S. LIS faculty". *Library & Information Research,* 28: 374-389.

Anuradha, K. T. and Urs, S. 2007. "Bibliometric indicators of Indian research collaboration patterns: A correspondence analysis." *Scientometrics,* 71(2): 179-189.

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

Dore, J. C. and Ojasoo, T. 2001. "How to analyze publication time trends by correspondence factor analysis: Analysis of publications by 48 countries in 18 disciplines over 12 years." *J. of the American Society for Information Science and Technology,* 52(9): 763-769.

Dore, J. C. et al. 1996. "Correspondence factor analysis of the publication patterns of 48 countries over the period 1981-1992." *J. of the American Society for Information Science,* 47(8): 588-602.

Greenacre, M. J. 1993. *Correspondence Analysis in Practice.* London: Academic Press.

Imperial, J. and Uez-Navarro, R.R. 2007. "Usefulness of Hirsch's *h*-index to evaluate scientific research in Spain." *Scientometrics,* 71(2): 272-282.

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

Van Raan, A. F. 2006. "Comparison of the Hirsch-index with standard bibliometric indicators and with peer judgment for 147chemistry research groups." *Scientometrics,* 67(3): 491-502.

〈Appedix 1〉 Frequency distribution of subject areas based on research productivity

Subject areas		Illinois	N Carolina	Syracuse	Washington	Michigan	Rutgers	Indiana	Pittsburgh	Texas	Florida State	total
HCI	frequency	1	8	3	6	0	3	1	4	5	2	33
	external strength (%)	3.0	24.2	9.1	18.2	0.0	9.1	3.0	12.1	15.2	6.1	100.0
	internal strenght (%)	2.4	9.8	4.1	7.8	0.0	4.9	1.7	6.9	15.2	3.3	5.7
ISB	frequency	7	4	5	15	8	9	2	7	5	1	63
	external strength (%)	11.1	6.3	7.9	23.8	12.7	14.3	3.2	11.1	7.9	1.6	100.0
	internal strenght (%)	16.7	4.9	6.8	19.5	23.5	14.8	3.4	12.1	15.2	1.7	10.9
IR	frequency	2	18	7	5	10	15	2	21	4	1	85
	external strength (%)	2.4	21.2	8.2	5.9	11.8	17.6	2.4	24.7	4.7	1.2	100.0
	internal strenght (%)	4.8	22.0	9.5	6.5	29.4	24.6	3.4	36.2	12.1	1.7	14.7
Info Sys	frequency	0	2	1	2	3	7	2	1	0	6	24
	external strength (%)	0.0	8.3	4.2	8.3	12.5	29.2	8.3	4.2	0.0	25.0	100.0
	internal strenght (%)	0.0	2.4	1.4	2.6	8.8	11.5	3.4	1.7	0.0	10.0	4.1
Mining_NLP	frequency	2	1	3	0	0	3	1	0	0	0	10
	external strength (%)	20.0	10.0	30.0	0.0	0.0	30.0	10.0	0.0	0.0	0.0	100.0
	internal strenght (%)	4.8	1.2	4.1	0.0	0.0	4.9	1.7	0.0	0.0	0.0	1.7
DL	frequency	1	3	1	1	0	4	0	3	4	2	19
	external strength (%)	5.3	15.8	5.3	5.3	0.0	21.1	0.0	15.8	21.1	10.5	100.0
	internal strenght (%)	2.4	3.7	1.4	1.3	0.0	6.6	0.0	5.2	12.1	3.3	3.3
Infometrics	frequency	4	2	4	1	0	0	18	7	1	4	41
	external strength (%)	9.8	4.9	9.8	2.4	0.0	0.0	43.9	17.1	2.4	9.8	100.0
	internal strenght (%)	9.5	2.4	5.4	1.3	0.0	0.0	30.5	12.1	3.0	6.7	7.1
Collabo	frequency	3	3	1	4	0	0	2	1	0	2	16
	external strength (%)	18.8	18.8	6.3	25.0	0.0	0.0	12.5	6.3	0.0	12.5	100.0
	internal strenght (%)	7.1	3.7	1.4	5.2	0.0	0.0	3.4	1.7	0.0	3.3	2.8
Organize Info	frequency	2	10	7	9	1	1	3	1	2	2	38
	external strength (%)	5.3	26.3	18.4	23.7	2.6	2.6	7.9	2.6	5.3	5.3	100.0
	internal strenght (%)	4.8	12.2	9.5	11.7	2.9	1.6	5.1	1.7	6.1	3.3	6.6
Info Manage	frequency	2	1	3	2	0	7	1	1	0	1	18
	external strength (%)	11.1	5.6	16.7	11.1	0.0	38.9	5.6	5.6	0.0	5.6	100.0
	internal strenght (%)	4.8	1.2	4.1	2.6	0.0	11.5	1.7	1.7	0.0	1.7	3.1
Archives	frequency	0	2	0	0	1	0	0	1	3	0	7
	external strength (%)	0.0	28.6	0.0	0.0	14.3	0.0	0.0	14.3	42.9	0.0	100.0
	internal strenght (%)	0.0	2.4	0.0	0.0	2.9	0.0	0.0	1.7	9.1	0.0	1.2
Manage	frequency	2	5	5	2	1	4	1	1	0	2	23
	external strength (%)	8.7	21.7	21.7	8.7	4.3	17.4	4.3	4.3	0.0	8.7	100.0
	internal strenght (%)	4.8	6.1	6.8	2.6	2.9	6.6	1.7	1.7	0.0	3.3	4.0
Ref_Info svc	frequency	3	9	12	9	0	2	3	0	3	5	46
	external strength (%)	6.5	19.6	26.1	19.6	0.0	4.3	6.5	0.0	6.5	10.9	100.0
	internal strenght (%)	7.1	11.0	16.2	11.7	0.0	3.3	5.1	0.0	9.1	8.3	7.9
Public Lib	frequency	0	1	0	1	1	0	0	0	0	6	9
	external strength (%)	0.0	11.1	0.0	11.1	11.1	0.0	0.0	0.0	0.0	66.7	100.0
	internal strenght (%)	0.0	1.2	0.0	1.3	2.9	0.0	0.0	0.0	0.0	10.0	1.6
Sch Lib_Child Lit	frequency	0	0	0	1	0	0	0	0	0	1	2
	external strength (%)	0.0	0.0	0.0	50.0	0.0	0.0	0.0	0.0	0.0	50.0	100.0
	internal strenght (%)	0.0	0.0	0.0	1.3	0.0	0.0	0.0	0.0	0.0	1.7	0.3
Info Policy	frequency	3	2	5	0	0	0	1	0	0	5	16
	external strength (%)	18.8	12.5	31.3	0.0	0.0	0.0	6.3	0.0	0.0	31.3	100.0
	internal strenght (%)	7.1	2.4	6.8	0.0	0.0	0.0	1.7	0.0	0.0	8.3	2.8
Telecomm_IT	frequency	0	1	8	4	2	2	2	1	2	3	25
	external strength (%)	0.0	4.0	32.0	16.0	8.0	8.0	8.0	4.0	8.0	12.0	100.0
	internal strenght (%)	0.0	1.2	10.8	5.2	5.9	3.3	3.4	1.7	6.1	5.0	4.3
Social Info	frequency	0	2	3	5	2	0	3	1	0	2	18
	external strength (%)	0.0	11.1	16.7	27.8	11.1	0.0	16.7	5.6	0.0	11.1	100.0
	internal strenght (%)	0.0	2.4	4.1	6.5	5.9	0.0	5.1	1.7	0.0	3.3	3.1
Special Info	frequency	1	6	2	0	0	0	2	1	1	1	14
	external strength (%)	7.1	42.9	14.3	0.0	0.0	0.0	14.3	7.1	7.1	7.1	100.0
	internal strenght (%)	2.4	7.3	2.7	0.0	0.0	0.0	3.4	1.7	3.0	1.7	2.4
Info Literacy	frequency	4	0	1	4	1	2	1	2	0	2	17
	external strength (%)	23.5	0.0	5.9	23.5	5.9	11.8	5.9	11.8	0.0	11.8	100.0
	internal strenght (%)	9.5	0.0	1.4	5.2	2.9	3.3	1.7	3.4	0.0	3.3	2.9
CMC	frequency	1	0	0	0	3	0	6	0	0	4	14
	external strength (%)	7.1	0.0	0.0	0.0	21.4	0.0	42.9	0.0	0.0	28.6	100.0
	internal strenght (%)	2.4	0.0	0.0	0.0	8.8	0.0	10.2	0.0	0.0	6.7	2.4
Compu assist Edu	frequency	1	0	1	1	0	0	0	2	1	5	11
	external strength (%)	9.1	0.0	9.1	9.1	0.0	0.0	0.0	18.2	9.1	45.5	100.0
	internal strenght (%)	2.4	0.0	1.4	1.3	0.0	0.0	0.0	3.4	3.0	8.3	1.9
LIS general	frequency	3	2	2	5	1	2	8	3	2	3	31
	external strength (%)	9.7	6.5	6.5	16.1	3.2	6.5	25.8	9.7	6.5	9.7	100.0
	internal strenght (%)	7.1	2.4	2.7	6.5	2.9	3.3	13.6	5.2	6.1	5.0	5.3
total	frequency	42	82	74	77	34	61	59	58	33	60	580
	external strength (%)	7.2	14.1	12.8	13.3	5.9	10.5	10.2	10.0	5.7	10.3	100.0
	internal strenght (%)	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

〈Appedix 2〉 Frequency distribution of subject areas based on research interests

Subject areas		Illinois	N Carolina	Syracuse	Washington	Michigan	Rutgers	Indiana	Pittsburgh	Texas	Florida State	total
HCI	frequency	6	4	5	6	7	4	2	9	6	2	51
	external strength (%)	11.8	7.8	9.8	11.8	13.7	7.8	3.9	17.6	11.8	3.9	100.0
	internal strenght (%)	7.7	4.7	4.9	9.2	10.9	9.1	3.4	10.1	10.2	3.2	7.2
ISB	frequency	1	4	1	7	1	4	1	4	4	4	31
	external strength (%)	3.2	12.9	3.2	22.6	3.2	12.9	3.2	12.9	12.9	12.9	100.0
	internal strenght (%)	1.3	4.7	1.0	10.8	1.6	9.1	1.7	4.5	6.8	6.3	4.4
IR	frequency	1	7	3	4	2	6	3	5	1	0	32
	external strength (%)	3.1	21.9	9.4	12.5	6.3	18.8	9.4	15.6	3.1	0.0	100.0
	internal strenght (%)	1.3	8.2	2.9	6.2	3.1	13.6	5.1	5.6	1.7	0.0	4.5
Info Sys	frequency	6	6	5	8	7	4	0	12	5	1	54
	external strength (%)	11.1	11.1	9.3	14.8	13.0	7.4	0.0	22.2	9.3	1.9	100.0
	internal strenght (%)	7.7	7.1	4.9	12.3	10.9	9.1	0.0	13.5	8.5	1.6	7.6
Mining_NLP	frequency	1	2	2	0	2	1	1	3	1	0	13
	external strength (%)	7.7	15.4	15.4	0.0	15.4	7.7	7.7	23.1	7.7	0.0	100.0
	internal strenght (%)	1.3	2.4	1.9	0.0	3.1	2.3	1.7	3.4	1.7	0.0	1.8
DL	frequency	4	3	2	0	0	3	2	4	5	4	27
	external strength (%)	14.8	11.1	7.4	0.0	0.0	11.1	7.4	14.8	18.5	14.8	100.0
	internal strenght (%)	5.1	3.5	1.9	0.0	0.0	6.8	3.4	4.5	8.5	6.3	3.8
Infometrics	frequency	2	2	3	1	0	2	9	4	0	3	26
	external strength (%)	7.7	7.7	11.5	3.8	0.0	7.7	34.6	15.4	0.0	11.5	100.0
	internal strenght (%)	2.6	2.4	2.9	1.5	0.0	4.5	15.3	4.5	0.0	4.8	3.7
Collabo	frequency	6	1	5	1	4	0	1	2	1	2	23
	external strength (%)	26.1	4.3	21.7	4.3	17.4	0.0	4.3	8.7	4.3	8.7	100.0
	internal strenght (%)	7.7	1.2	4.9	1.5	6.3	0.0	1.7	2.2	1.7	3.2	3.2
Organize Info	frequency	6	8	5	5	3	1	5	7	3	3	46
	external strength (%)	13.0	17.4	10.9	10.9	6.5	2.2	10.9	15.2	6.5	6.5	100.0
	internal strenght (%)	7.7	9.4	4.9	7.7	4.7	2.3	8.5	7.9	5.1	4.8	6.5
Info Manage	frequency	3	4	10	4	2	0	2	3	3	4	35
	external strength (%)	8.6	11.4	28.6	11.4	5.7	0.0	5.7	8.6	8.6	11.4	100.0
	internal strenght (%)	3.8	4.7	9.7	6.2	3.1	0.0	3.4	3.4	5.1	6.3	4.9
Archives	frequency	2	2	1	0	5	2	1	1	4	1	19
	external strength (%)	10.5	10.5	5.3	0.0	26.3	10.5	5.3	5.3	21.1	5.3	100.0
	internal strenght (%)	2.6	2.4	1.0	0.0	7.8	4.5	1.7	1.1	6.8	1.6	2.7
Manage	frequency	2	8	6	3	1	0	2	2	2	6	32
	external strength (%)	6.3	25.0	18.8	9.4	3.1	0.0	6.3	6.3	6.3	18.8	100.0
	internal strenght (%)	2.6	9.4	5.8	4.6	1.6	0.0	3.4	2.2	3.4	9.5	4.5
Ref_info svc	frequency	2	5	5	3	0	2	3	2	2	4	28
	external strength (%)	7.1	17.9	17.9	10.7	0.0	7.1	10.7	7.1	7.1	14.3	100.0
	internal strenght (%)	2.6	5.9	4.9	4.6	0.0	4.5	5.1	2.2	3.4	6.3	3.9
Public Lib	frequency	1	1	0	0	0	0	0	0	1	2	5
	external strength (%)	20.0	20.0	0.0	0.0	0.0	0.0	0.0	0.0	20.0	40.0	100.0
	internal strenght (%)	1.3	1.2	0.0	0.0	0.0	0.0	0.0	0.0	1.7	3.2	0.7
Sch Lib_Child L	frequency	4	4	1	0	0	2	0	1	2	3	17
	external strength (%)	23.5	23.5	5.9	0.0	0.0	11.8	0.0	5.9	11.8	17.6	100.0
	internal strenght (%)	5.1	4.7	1.0	0.0	0.0	4.5	0.0	1.1	3.4	4.8	2.4
Info Policy	frequency	3	3	9	4	6	0	3	4	1	2	35
	external strength (%)	8.6	8.6	25.7	11.4	17.1	0.0	8.6	11.4	2.9	5.7	100.0
	internal strenght (%)	3.8	3.5	8.7	6.2	9.4	0.0	5.1	4.5	1.7	3.2	4.9
Telecomm_IT	frequency	4	0	14	1	5	1	1	4	0	3	33
	external strength (%)	12.1	0.0	42.4	3.0	15.2	3.0	3.0	12.1	0.0	9.1	100.0
	internal strenght (%)	5.1	0.0	13.6	1.5	7.8	2.3	1.7	4.5	0.0	4.8	4.7
Social Info	frequency	11	5	11	7	11	2	6	6	7	5	71
	external strength (%)	15.5	7.0	15.5	9.9	15.5	2.8	8.5	8.5	9.9	7.0	100.0
	internal strenght (%)	14.1	5.9	10.7	10.8	17.2	4.5	10.2	6.7	11.9	7.9	10.0
Special Info	frequency	1	5	0	3	4	1	1	3	2	4	24
	external strength (%)	4.2	20.8	0.0	12.5	16.7	4.2	4.2	12.5	8.3	16.7	100.0
	internal strenght (%)	1.3	5.9	0.0	4.6	6.3	2.3	1.7	3.4	3.4	6.3	3.4
Info Literacy	frequency	2	0	2	2	1	2	2	1	0	2	14
	external strength (%)	14.3	0.0	14.3	14.3	7.1	14.3	14.3	7.1	0.0	14.3	100.0
	internal strenght (%)	2.6	0.0	1.9	3.1	1.6	4.5	3.4	1.1	0.0	3.2	2.0
CMC	frequency	1	0	0	0	0	0	4	3	0	3	11
	external strength (%)	9.1	0.0	0.0	0.0	0.0	0.0	36.4	27.3	0.0	27.3	100.0
	internal strenght (%)	1.3	0.0	0.0	0.0	0.0	0.0	6.8	3.4	0.0	4.8	1.6
Compu assist E	frequency	3	3	5	1	1	2	0	4	1	0	20
	external strength (%)	15.0	15.0	25.0	5.0	5.0	10.0	0.0	20.0	5.0	0.0	100.0
	internal strenght (%)	3.8	3.5	4.9	1.5	1.6	4.5	0.0	4.5	1.7	0.0	2.8
LIS general	frequency	6	8	8	5	2	5	10	5	8	5	62
	external strength (%)	9.7	12.9	12.9	8.1	3.2	8.1	16.1	8.1	12.9	8.1	100.0
	internal strenght (%)	7.7	9.4	7.8	7.7	3.1	11.4	16.9	5.6	13.6	7.9	8.7
total	frequency	78	85	103	65	64	44	59	89	59	63	709
	external strength (%)	11.0	12.0	14.5	9.2	9.0	6.2	8.3	12.6	8.3	8.9	100.0
	internal strenght (%)	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
number of researchers		26	24	31	26	38	17	25	33	18	24	262
average research interests		3.0	3.5	3.3	2.5	1.7	2.6	2.4	2.7	3.3	2.6	2.7