

From the United States to China: the transfer of research centres in Information Science

Jianhua Hou^{1,*} and Xiucui Yang²

¹School of Information Management, Sun Yat-sen University, Waihuan East Road, Guangzhou, Guangdong 510006, China

²College of Economics and Management, Dalian University, Dalian 116622 and Dalian Economic Technological Development Zone, Dalian 116622, China

This study not only analyses the centres of research in Information Science (IS), including the migration of central topics and central countries, but also analyses the relationship between the shifting of centres of research and their transformation. In addition, this study explores the relationship between the formation of the centre of research and the academic influence of the country on IS itself. We collected 25,150 articles, including 313,293 references about citation analysis, from databases SCI-E and SSCI between 1977 and 2016 as our data source. The following findings were obtained through this study: the transfer (transfer time) of central research topics in the IS domain has accelerated, from 12 to 8 years between 1980 and 1990, to 6 to 4 years between 2000 and 2010, and to 3 years between 2011 and 2016. The number of central research topics has also grown, from one between 1997 and 2006, to two from 2006 to 2013 to three from 2013 to 2016. The geographical centres of IS research were the US and Britain between the 1970s and 1980s, but gradually migrated through neighbouring countries, and finally to Asia by 2000. China, which became the centre of research for IS in 2005 for the first time, has been ranked first since 2011. In addition, countries acting as centres of research enjoy not only a high output of literature but also great academic influence. The theoretical and practical implications of our findings are discussed.

Keywords: CiteSpace software, Information Science, scientometric analysis, transfer of research centre.

EXPLORING the rules in the development of science is an important topic in scientometrics. Understanding the objective laws of the development of science can be undertaken from two aspects, namely, the development rule of science itself and the relationship between scientific development and the external environment. In terms of the former aspect, Fremont Ryder, an American librarian, and Derek John de Solla Price, a scientometrician, founded the rule of accelerated growth of scientific literature in the 1940s. Since then, other researchers have also proved that the growth in the number of scientific

achievements, disciplines, scientific research funds and researchers are all increasing. Between the 1960s and 1970s, Mintomo Yuasa, a Japanese historian of science, and Zhao Hong-zhou, a Chinese scientometrician, revealed the transfer law of the centre of scientific activity and calculated the period of scientific prosperity of each country (namely, their period of being a centre of research) respectively, and independently through statistical analyses of major scientific achievements. This follow-up study proves the relationship of disciplines with regard to the shifting time and the geographical distribution of centres of scientific activity. With rapid development and growing complexity of science, research topics in different fields are changing at an accelerating rate, and the number of research institutions is increasing continually. Then, does the migration of centres of research in Information Science (IS) also have its own rule? What is the relationship between the change in topics of research and the migration of its geographical centre?

The era of big data has had a great impact on conventional research in IS. With rapid development in information technology and internet technology, in particular, IS shows a tendency toward new development and rapid changes (such as new paradigms, objectives, methods of research and new analysis tools). Computer technology plays an important role in more than two-thirds of the research fields in IS¹. This reveals a significant knowledge flow from computer science to information science². The research objective of information science has shifted from one main source, namely, paper publications, to multiple information sources including electronic texts, internet data, physical data, video and audio.

Based on traditional mathematical statistics, mathematical modelling, and classical literature laws, diverse research methods in relation to computer-aided analysis have been developed. For example, content analysis, experiments, and theoretical approaches are extensively applied in this field³. The research level has also shifted from the analysis of information retrieval and text structure to in-depth analysis, including analysis of contextual content and latent semantics⁴. In recent years, the amount of research literature on IS has increased rapidly, and new trends and front topics of research are

*For correspondence. (e-mail: houjh5@mail.sysu.edu.cn)

constantly emerging. Moreover, the distribution of institutions and countries producing this research is changing rapidly.

This study conducts an in-depth and comprehensive analysis that is not only a visualization of the evolution of research topics in IS but also a comparative analysis of central topics and central countries by answering the following: (1) What is the relationship between changes in the central research topic and the country as centre of research? (2) What is the relationship between the country as centre of research and its academic influence on IS?

Related work

Transfer of active science centre

W. C. Dampier, a British historian of science, put forward the idea of a 'world science centre' in his work 'A History of Science and its Relations with Philosophy and Religion'. In 1954, Bernard introduced the idea of a 'centre of technological and scientific activity' in his book, *Science in History*, and enumerated the laws of the transfer of centres of technological and scientific activity in history. Inspired by this book, Yuasa⁵ discovered the rule of the transfer of centres of scientific activities: a country whose scientific achievements account for over 25% of the total achievements in the world in the same period can be defined as a centre of scientific activity. Scientific prosperity means the period in which a country is a world centre of scientific activity.

Zhao⁶ also discovered this phenomenon. In 1985, he pointed out the optimum age for making scientific discoveries through statistical methods, and formulated an experiential formula involving the number of scientific achievements, the number of scientists, and their respective ages, and explained the Yuasa Phenomenon to some extent⁷. Liang *et al.*⁸ explored the geographical and chronological characteristics of shifts in the world's centre of scientific activity on the basis of 4087 records retrieved from two chronological tables about the history of science and technology. By comparing the shifting times of the world's discipline centres and science centres, they revealed the relationship between leading disciplines in countries (including Italy, UK and USA) and 'excavation-worthy' disciplines. They also explored the internal mechanism of shifts in world science centres.

With the accelerating development of scientific research and cognitive activities, cross-region and cross-nation scientific collaboration is increasingly popular, and collaborative research is generally cited more often than research by a single author⁹. With the increasingly blurred boundaries of the centres of scientific activity, it is obvious that many countries may become the world's

centre of scientific activity jointly. Disciplines constituting the complex knowledge system of science, research topics, and geographical distribution of the field are also changing and transforming rapidly.

According to Bernard, who put forward the concept of a 'centre of technological and scientific activity', this research proposes a concept related to research centres in the IS discipline that includes a central research topic and central research location, that is, the institution or country that generates the central research topic. Yuasa and Hong-Zhou Zhao introduced the quantitative standard for research centre certification: to qualify as a research centre, its scientific achievements must account for more than 25% of the world's total scientific achievement during the same period of time. This standard is widely accepted and recognized.

However, previous scientometric and visualization analysis of the IS field mainly focused on the research topics, the research frontiers and the evolution of IS through citation analysis method (DCA or ACA), and did not reveal the relationship between central research topics and geographical locations (or institutions). Additionally, the research also did not mention the relationship between central research topic, and the academic influence of the geographical location.

Therefore, the purpose of this study is to know the relationship between the change of central research topic and the migration of the geographical centre of research in IS through scientometric analysis and visualization of IS research topics during the years 1977–2016. Additionally, the study also reveals the relationship between the country as centre of research and the overall academic influence of these countries in IS.

Therefore, as shown by Yuasa⁵ and Zhao⁷, this study also uses 25% as the selection criteria. Then, the top 25% clusters of research topics in one year are the central research topics of that year. The size of a cluster can be calculated through the amount of literature in each cluster. A research topic whose size accounts for over 25% of the total cluster for one year can be defined as a central research topic for that year. Then, we extract the distribution of institutions and countries that produced the most literature as geographical research centres.

Scientometric analysis in IS field

Science mapping based on scientometric methods is also an important topic of research in the fields of scientometrics and IS. Researchers study the knowledge structure and dynamic evolution of IS through related scientific literature of IS. The number of related research articles on the evolution of research topics and research fronts has increased rapidly. The data analysed in the literature above was mainly derived from 12 core journals. The previous author co-citation analyses used these

RESEARCH ARTICLES

Table 1. Related research of information sciences using 12 core journals

Study	Period	Data	Analysis	Topics or sub-field
White and McCain ¹¹	1972–1995	12 core journals	ACA	Experimental retrieval, practical retrieval, citation analysis/bibliometrics, citation theory, scientific/scholarly communication, user theory
White ¹³	1972–1995	12 core journals	Pathfinder networks	
Zhao and Strotmann ¹⁴	1996–2005	12 core journals	ACA, FA	User theory, evaluative citation analysis, experimental retrieval, Webometrics, Science communication, visualization of knowledge domains, user's judgements of relevance webometrics, image retrieval, web searching
Zhao and Strotmann ¹⁵	1996–2005	12 core journals	ABCA	
Zhao and Strotmann ¹⁶	1996–2005	12 core journals	All-author-based ACA	
Ibekwe-SanJuan ¹⁷	1996–2008	12 core journal	Text mining	Automated information retrieval, web-based studies, citation studies, vector space, open access, Google scholar and h-index
Klavans ¹⁸	2001–2005	12 core journal	Coco analysis	Citation analysis, information retrieval, information behaviour
Chen <i>et al.</i> ⁴	1996–2008	12 core journals	Multiple co-citation	H-index, Interactive Information retrieval, academic web, Information retrieval, citation behaviour
Klavans and Boyack ⁴²	2000–2008	12 core journals	DCA, local and global map	Information seeking behaviour, scientometrics, co-citation analysis, citation behaviour, computer-enhanced retrieval
Zhao and Strotmann ¹⁰	2006–2010	12 core journals	ACA, ABCA	H-index, information behaviour, mapping of science, webometrics, relevance, IR systems, IS theories and foundations
Yang <i>et al.</i> ¹⁹	2006–2015	12 core journals	AKCA	IR, patent analysis, open access, mapping of science, bibliometric evaluation

journals to define the IS research field and are expected to be updated in present studies (Table 1)¹⁰.

The main research methods include document co-citation analysis (DCA), author co-citation analysis (ACA), co-word analysis and cluster analysis. Information visualization and science mapping tools can be used to explore issues such as the evolution of research topics, research fronts and hotspots of IS from different perspectives. White and McCain¹¹ were the first to analyse the research topics of IS through 12 core journals. They analysed research topics between 1972 and 1995 through ACA. Subsequent studies^{10,12–19} revealed the research topics of IS in different periods. This study applied different citation analysis methods to explore the research topics and knowledge structures of IS, conducted science mapping for different periods, and manually identified the research topics and knowledge structures of IS. This study claims that the IS domain can be divided into two camps: information retrieval/seeking and citations/bibliometrics. In recent years, some studies have also explored new research topics such as the H-index, information behaviour, science mapping, web metrics, patent analysis, open access and bibliometric evaluation^{10,19}.

Some researchers chose different journals to define the research topic of IS. White and Griffith²⁰ put forward ACA, mapped and visualized an authors group with 39 representative authors, and revealed the knowledge structure of IS with five main author groups. Adopting litera-

ture published in five IS journals between 2002 and 2004, Janssens *et al.*^{21,22} analysed research fronts of IS through a full-text analysis. There were four clusters in total with different topics including bibliometric research, information retrieval research, web metrics and patent analysis, the last two being new clusters. Selecting eight IS core journals through journal co-citations, Besselaar and Heimeriks²³ put forward a co-word analysis based on JASIST with a method of word-reference co-occurrences to analyse the knowledge structure of IS between 1986 and 2002.

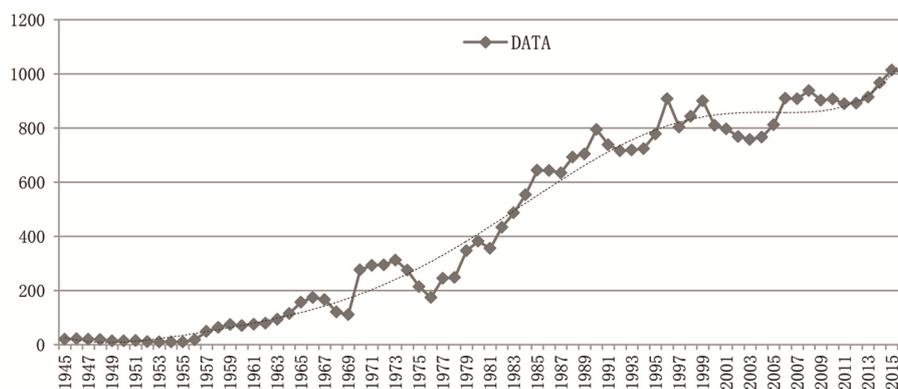
Astrom¹² took 21 IS journals as his data source and detected the research front topics of IS between 1990 and 2004, including information searches and information retrieval. Milojević *et al.*¹⁸ took IS journals as a dataset and revealed the cognitive structure of library and IS by composing a suite of analyses of words in article titles. The result showed that IS has three main branches: libraries, information and science. Some researchers chose JASIST, the most representative journal in the IS domain, as a database, to analyse the knowledge structure of IS²⁵. Chang and Huang²⁶ adopted 580 highly cited studies of IS and found that the research fronts of IS between 1995 and 2014 mainly included information retrieval, web searches and bibliometrics.

Lei and Yan²⁷ analysed the readability of abstracts and full texts of articles published in four journals of IS from 2003 to 2012. Khasseh *et al.*²⁸ studied the intellectual

Table 2. List of the 12 core journals recorded by Web of Science

Journal title	Period	Amount of literature
<i>ARIST</i>	1966–2011	550
<i>Electronic Library</i>	1984–2016	2907
<i>Information Processing and Management</i>	1963–2016	3603
<i>Information Technology and Libraries</i>	1982–2016	1670
<i>Journal of Documentation</i>	1945–2016	3986
<i>Journal of Information Science</i>	1979–2016	2059
<i>JASIS/JASIST</i>	1970–2016	6200
<i>Library and Information Science Research</i>	1984–2016	1349
<i>Library Resources and Technical Services</i>	1957–2016	3069
<i>Program-Automated Library and Information Systems</i>	1979–1996	995
<i>Program-Electronic Library and Information Systems</i>	1981–2016	1504
<i>Scientometrics</i>	1978–2016	4747

Note. The journal *ARIST* was started in 1963 and stopped in 2012. The journal *Program-Automated Library and Information Systems* was started in 1979 and stopped in 1997. *Information Processing and Management* was started in 1963, renamed *Information Storage and Retrieval* during 1963–1974, and changed back to the original name in 1975. *JASIS/JASIST* started in 1950 was included in the database in 1970. It was renamed *Journal of the American Society for Information Science* between 1970 and 2000, *Journal of the American Society for Information Science and Technology* in 2011 and *Journal of the Association for Information Science and Technology* in 2014.

**Figure 1.** Change in the total number of articles published in the 12 core journals.

structure of iMetrics during the period 1978 to 2014 using data from *Scientometrics*, the *Journal of Informetrics* and relevant articles in six journals. This study took documents from different specified periods as the overall object to analyse the research topics and research fronts of information science in different periods. The researchers proved that the main topics and subject areas of IS have been evolving¹⁰.

Data and method

Data collection

From the existing research, majority of metrological analyses of the definition, knowledge structures, and research topics of IS used 12 core journals of IS as the data source, in spite of some deficiencies²⁹. However, those researchers made long-term tracking studies of IS with data retrieved from these journals and through similar

methods¹⁰. Since journals are dynamic and changing, there is no constant ranking of journal influence. In order to maintain the relative consistency and continuity of data samples, the sample data were selected from the representative 12 core journals that had been widely recognized for a relatively long time. This paper also takes these 12 journals as data sources for the comparative analysis of the evolution of IS.

The dataset used for the present study includes 12 core journals of IS recorded in the database from Web of Science (WoS; Table 2). A total of 32,639 records were retrieved from these journals starting from their first publication until 2016 (Figure 1). The resultant dataset includes 19,496 articles, accounting for 59.73% of the total references, and 7674 book reviews, accounting for 23.51%. The remaining 5469 records account for 16.76% of the total number of records that include literature related to Biography, Unspecified, News, Case Reports, Editorials and others.

Methods and tools

Although an increasing number of research studies employ ACA in identifying research topics of IS or other subfields, and there are many advanced methods such as author direction citation analysis (ADCA), author keyword coupling analysis (AKCA), and author bibliographic coupling analysis (ABCA), DCA is capable of producing more detailed results. Moreover, explaining DCA is easier than explaining ACA because DCA is more specific while ACA covers a larger range⁴. The knowledge structure of a research field can be detected through cluster analyses in DCA. Because of the mapping relation between the references and cited documents, the cluster results of document co-citation and the topic words of the cited documents also have a mapping relation. Therefore, the topics of the cited documents can be used to label academic communities and scientific knowledge structures of the research field displayed by DCA.

This study employs CiteSpace to extract topic words and map them to the document co-citation cluster, display them by visualizing knowledge mapping, and detect the research topics of IS year by year more precisely. Based on this mapping relation, the distribution information of the institutions and countries that produced these cited documents mapping the central research topics can be extracted, analysed by Google Fusion Table, and visualized on Google Maps. Google Fusion Tables display the distribution of these research institutions and countries, and also provide maps of hot regions among these countries and institutions.

CiteSpace is a software system developed by Chaomei Chen from Drexel University and is based on the Java programming language. It is an information visualization application used for multiplex, time-sliced, dynamic complex network analysis^{30,31}. CiteSpace is mainly applied to detect and analyse trends of changes in the research front of a knowledge domain, the relationship between the research front and its knowledge base, and the internal relations between different research fronts. Through a visualized analysis of IS, researchers are intuitively able to identify the path of evolution of research topics and the classical primary literature of the knowledge domain. This paper adopts CiteSpace to draw document co-citation network mapping in the field of IS year by year, and employs spectrum cluster analysis to conduct a cluster analysis of the co-citation network. A clustering algorithm has a distinct advantage in identifying the optimal number of clusters. Moreover, the cluster is set upon the basis of the topological structure and the link strength of DCA (ref. 4).

Google Fusion Tables is an online platform provided by Google for data management and visualization applications. It is convenient for saving, managing, co-editing, visualizing and sharing data tables online. It can also visualize the geographical distribution of big data³²⁻³⁴.

(<https://www.google.com/fusiontables/data?dsrcid=implicit>).

We apply CiteSpace first to extract the information of the institutions and nations that produced the cited articles on central research topics year by year, and then apply Google Fusion Tables to further visualize and analyse the geographical locations of this data. The steps are as follows: (1) Use CiteSpace to clean and remove duplicate data downloaded from the database. (2) After cleaning, input data into the layout of 'Generate Google Earth Maps (KML2.0)' under the Geographical menu linked to CiteSpace (Figure 2). (3) After Google Earth processes all of the data, input the comma-separated values files (.csv) in Google Fusion Tables year by year for visualization display and analysis (Figure 3). (4) Transform comma-separated values documents into Microsoft Office Excel (.xlsx) documents if any error message appears when analysing comma-separated values documents through Google Fusion Tables.

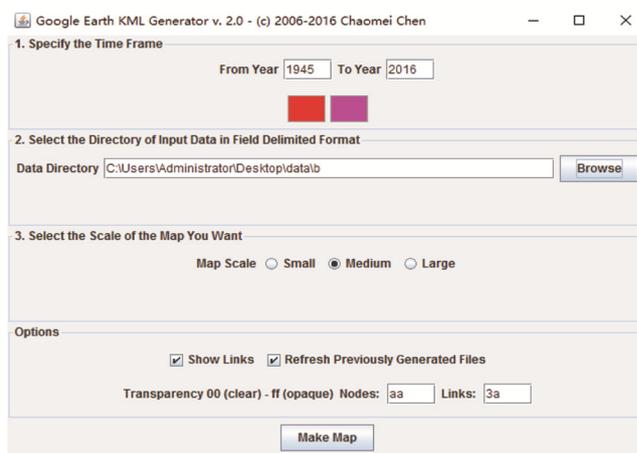


Figure 2. Setup interface to generate Google Earth maps.



Figure 3. Setup interface of Google Fusion tables.

Table 3. List of research topics in IS domain year by year and leading institutions in publishing these research topics

Year	Cluster identifier (size)	Central research topics	Main institutions
1977	Bibliographic data-base (15%) Information-processing standard (13%)	Information retrieval technique	University of Michigan, University of Ibadan, Aston University, University of London, University of Illinois
1978	Online system (24%) Library resource sharing (18%)		University of California, University of Wales, University of Alberta, Massachusetts Institute of Technology, University of Illinois
1979	Non-book materials (18%) Model-based online management information-system (17%)		University of Illinois, St John's University, Clarkson University, University of California, George Washington University
1980	Scientific journal (17%) Information technology (16%)		University of Sheffield, University of Pennsylvania, Institute for Scientific Information (Philadelphia, US), University of Western Ontario, University of Illinois
1981	Online information-retrieval system (6%) Probabilistic model (6%) Departmental library (6%) Boolean retrieval-systems (5%) Library network (4%)		Louisiana State University, Politechnika Wroclawska (Poland), Massachusetts Institute of Technology, University of Massachusetts, Osaka Prefecture University
1982	Subject analysis (29%)		University of Western Ontario, University of Toronto, University of Wisconsin, University of Copenhagen, University of Sheffield
1983	Budget preparation (16%) Scientific activities (14%)		Hungarian Academy of Sciences, University of North Carolina, City University of London, University of Bradford, University of California
1984	Quantitative studies (9%) Database management system (9%) Information-retrieval system-design (8%)		Drexel University, University of Southern California, Vanderbilt University, Louisiana State University, University of Illinois
1985	Personal information-systems (12%) Online public-access catalog (8%) Fuzzy-set applications (7%)		Institute for Scientific Information (Philadelphia, USA), Hungarian Academy of Sciences, Flushing University, Louisiana State University, Leiden University
1986	Retrieval performance (17%) Bradford law (17%)		University of California, City University of London, University College Dublin, Hungarian Academy of Sciences, Cornell University
1987	Retrieval technique (22%) Sampling properties (16%)		University of Western Ontario, Hungarian Academy of Sciences, Louisiana State University, University of North Carolina, University of Amsterdam
1988	Online retrieval systems (14%) Foreign-language barrier (12%)		The State University of New Jersey, University of Sheffield, University of Michigan, University of Western Ontario, Indiana University
1989	Scientometric study (11%) System-design (10%) Bradford distribution (9%)	Scientometric study	University of Sheffield, Drexel University, Leiden University, University of Western Ontario, Indian Institute of Chemical Biology
1990	Bradford law (15%) Knowledge organization (14%)		City University of London, University of California, Washington University, University of Western Ontario, University of Osnabrück
1991	User interface (17%) Database model (13%)	Information retrieval system	University of California, University of North Carolina, University of Maryland, University of South Carolina, University of Sheffield
1992	Bibliographic retrieval-system (15%) Frequency-distribution (14%)		University of Maryland, University of California, University of North Carolina,, University of Wales, Instituto de Información y Documentación en Ciencia y Tecnología (Spain)

(Contd)

RESEARCH ARTICLES

Table 3. (Contd)

Year	Cluster identifier (size)	Central research topics	Main institutions
1993	Information retrieval (18%) Technical service (14%)		The Catholic University of America, University of Massachusetts, University of North Carolina, Dalhousie University, Korea Advanced Institute of Science and Technology
1994	Document retrieval (11%) Retrieval system (10%) Informetric distribution (9%)		University of North Carolina, Indiana University, University of Michigan, University of California, Wroclaw University of Technology
1995	Displaying online catalog postings (16%) Scientific literature (15%)		Hungarian Academy of Sciences, Leiden University, University of Maryland, University of Tampere, Indiana University
1996	Information retrieval (11%) Document retrieval system (11%) Document retrieval (9%)		University of California, University of North Texas, University of Sheffield, Cornell University, University of Copenhagen
1997	Search interface (18%) Similarity measure (14%)		University of North Carolina, City University of London, University of Illinois, University of Chicago, University of West Ontario
1998	Searching bibliographic information retrieval systems (14%) Optimizing similarity (11%)		University of Arizona, University of Pittsburgh, Louisiana State University, Concordia University, University of North Carolina
1999	Author co-citation analysis (19%) Document retrieval (18%)	Author co-citation analysis	University of Sheffield, University of Illinois, Indiana University, Universidad de Granada, University of New South Wales
2000	User-centred interface (13%) Information retrieval (13%)	Information retrieval behaviour	Concordia University, University of Tennessee, McGill University, Loughborough University, Lancaster University
2001	Query expansion (13%) Critical view (12%)		University of North Carolina, Leiden University, Drexel University, University of Tampere, Hungarian Academy of Sciences
2002	Relevance judgement (15%) Relevance criteria (11%)		University of North Carolina, University of Tennessee, McGill University, City University of London, Drexel University
2003	Open source software development (16%) Electronic publishing (15%)		University of Wolverhampton, Indiana University, University of West Ontario, Royal Netherlands Academy of Arts and Sciences, University of Amsterdam
2004	Retrieval system (10%) Search engine coverage bias (8%) Critical view (7%)		University of Wolverhampton, University of West Ontario, University of Sheffield, The Hebrew University of Jerusalem, University of Arizona
2005	Research performance (14%) Retrieval system (12%)		Transnational University Limburg, University of Pennsylvania, McGill University, University of Wolverhampton, University of Amsterdam
2006	Knowledge management (13%) Bibliometric information (12%)	Information behaviour, H-index	University of Wolverhampton, University of Amsterdam, University Hasselt, Leiden University, KU Leuven
2007	Human information behaviour (15%) H-Index (12%)		University Hasselt, Indiana University, Rutgers University, University of Copenhagen, University of Amsterdam
2008	Information ethics (11%) Science system (9%) H-Index (8%)		University of Wolverhampton, University Hasselt, Swiss Federal Institute of Technology Zurich, KU Leuven, University of Amsterdam
2009	Information retrieval (13%) International scientific collaboration (12%)		University of Pennsylvania, University of Copenhagen, University of Sheffield, University of Tampere, University of Amsterdam

(Contd)

Table 3. (Contd)

Year	Cluster identifier (size)	Central research topics	Main institutions
2010	Bibliometric mapping (14%) Information literacy (12%)	Science knowledge	University of Amsterdam, Indiana University, University of London, University of Pennsylvania, Leiden University
2011	H-Index (23%) Bibliometric index (10%)	H-Index, citation network analysis	Indiana University, University of Amsterdam, Politecnico di Torino Polytechnic University of Turin, University of Hasselt, India Academy of Sciences
2012	Citation networks (13%) Semantic patent analysis (12%)		University of Amsterdam, Indiana University, University of Tokyo, Seoul National University, Chinese Academy of Sciences
2013	Probabilistic clustering (13%) Scholarly monograph (13%)		University of Amsterdam, Indiana University, Nanyang Technological University, University of Sheffield, Dalian University of Technology
2014	Scientometric indicator (14%) Social networking site (12%)	Scientific indicators, scientific collaboration, altmetrics	Max Planck Society, University of Granada, University of Copenhagen, Nanjing University, CSIR National Institute for Interdisciplinary Science and Technology (India)
2015	Scientific collaboration (13%) International collaboration (12%)		University of Wolverhampton, Drexel University, University of Chinese Academy of Sciences, Polytechnic University of Valencia, Max Planck Society
2016	Twitter count (14%) Nanotechnology research (13%)		Max Planck Society, Institute for System Analysis and Computer Science (National Research Council of Italy), University of Wolverhampton, DZHW - German Centre for Higher Education Research and Science Studies (Germany), Chinese Academy of Sciences

Note: From 1977 to 1979, only four countries were extracted from the literature of central topics. China, when mentioned here and in the following paragraphs, includes Hong Kong, Macao, Taiwan, and the United Kingdom includes Scotland, Northern Ireland and Wales.

Results

Change in the central research topics

This paper employs CiteSpace software, selects ‘Cited Reference’ under the ‘Node Types’ menu, selects ‘Top N’ under the ‘Selection Criteria’ menu, sets the threshold to Top 100, and conducts DCA and cluster analyses on literature between 1977 and 2016, year by year. The cluster results of each year are arranged from highest to cluster size (including the number of nodes), accounting for over 25% of the total research topics of that year (Table 3).

On the whole, the main research topics of IS still focus on two camps, namely, information retrieval and knowledge domains¹⁰, with the former being the mainline of research in this field. Between 1977 and 1988, the research topics were centred on information retrieval techniques including bibliographic databases, information processing standards, information technology, subject analysis, and retrieval techniques, with emphasis on the study and solution of problems about information retrieval techniques.

In 1989, the scientometric study became the main topic of information science for the first time. One reason is that scientometrics developed rapidly from the 1970s to the 1980s, within scientometrics analysis of document in-

formation in an important part³⁵. Another reason is that scientometrics methods have been extensively used in information processing technology such as information retrieval. Between 1991 and 1998, the main research topic was information retrieval systems, including document retrieval systems, online catalog postings, similarity measures and optimization. Research topics in this period started to transfer from information retrieval techniques in the last period to the systematization of information retrieval.

Since the 1990s, with the rapid and explosive development of information, searching useful information accurately and efficiently has attracted the most attention in the field of IS. In 1999, author co-citation analysis became the main research topic for the first time. Author co-citation analysis was first put forward by White and Griffith²⁰. White and McCain¹¹ studied the knowledge structure of IS through an author co-citation analysis based on 12 IS core journals, and exerted great influence on the follow-up methodology of IS.

After 2000, the central research topics of IS changed even faster, and multiple central research topics emerged during the same period. The central research topic between 2000 and 2005 was information retrieval behaviour, including user-centred interfaces, query expansion and relevance judgment. Because of the great influence of

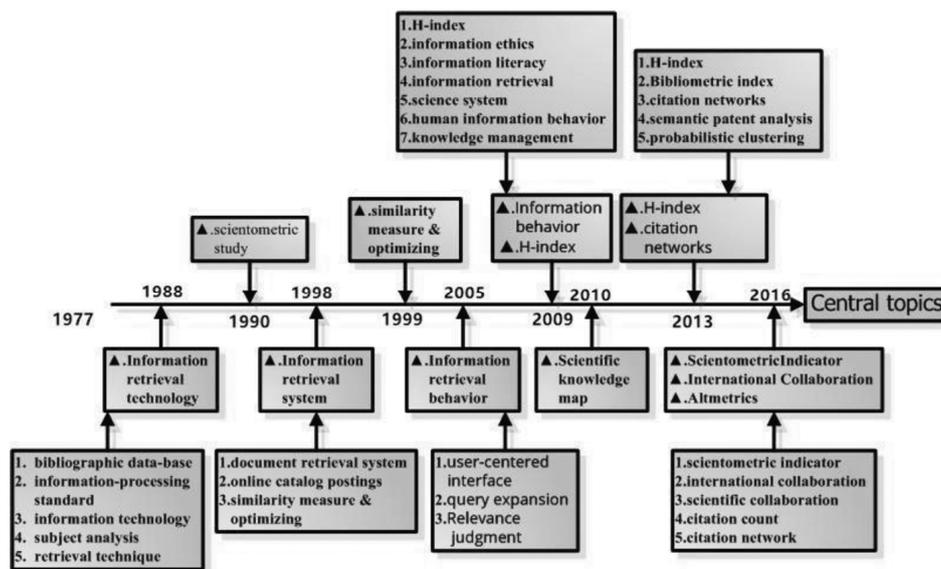


Figure 4. Evolution map of central research topics.

information retrieval techniques on users and society, researchers inevitably paid more attention to the development of information retrieval techniques. Information retrieval behaviour not only developed in the field of technology, but was also integrated and developed in culture and society.

Between 2006 and 2009, there were two central research topics: (1) information behaviour (human information behaviour, information ethics, information literacy, and information retrieval), and (2) H-Index (H-index, science system and knowledge management). In 2006, with the further development of information and network technology, information behaviour, information ethics, and information ability became the focus of researchers. Meanwhile, the H-index, as the quantitative research on science, became the central research topic during this period. Hirsch³⁶ put forward the H-index for the first time, and exerted a great influence on research related to IS and scientometrics. As a scientific quantitative evaluation index, the H-index has been applied and developed rapidly. In 2007, the H-index became a central research topic of IS and ranked first among all research topics in 2011.

In 2010, science mapping became the central research topic. With the development of information and computer technology, the citation relations in literature could be mapped into visualized knowledge networks, and the efficiency of analyses of scientific knowledge structure was also greatly improved. Information and computer technology was vital in related analyses of information retrieval and knowledge domains.

In 2011, the central research topic of IS shifted to the study of technology and methods. Two central research topics in the period 2011–2013 were (1) the H-Index (H-index and bibliometric index) and (2) citation network analysis (citation networks, semantic patent analysis and

probabilistic clustering). Three central research topics emerged between 2014 and 2016: (1) scientometric indicators (scientometric indicators and social networking sites), (2) scientific collaboration (scientific collaboration and international collaboration), and (3) altmetrics (Twitter count) (Figure 4)^{37–41}.

Geographical migration of centres of research

Central research institutions are sampled according to the number of annual publications that are related to the central topic. The top five are selected as the central research institutions in each of the sample periods, and are listed in Table 3 from high to low. In addition, at each time point, the number of publications of the top-five central research institutions that are related to the central topic must exceed 25% of the total number of publications that are related to the central topic during the same period of time.

Similarly, research centre countries are selected according to the number of annual publications that are related to the central topic. The top five are selected as the central research countries in each of the sample periods and are listed in Figure 5 from high to low. In addition, at each time point, the number of publications of the top-five research centre countries that are related to the central topic must exceed 25% of the total number of publications that are related to the central topic during the same period of time.

Research centre countries and research centre institutions may or may not be correlated. There are three major situations when research centre countries are on the list: (1) A country has developed few institutions engaging in related research, but these few institutions have many publications, resulting in a high overall publication

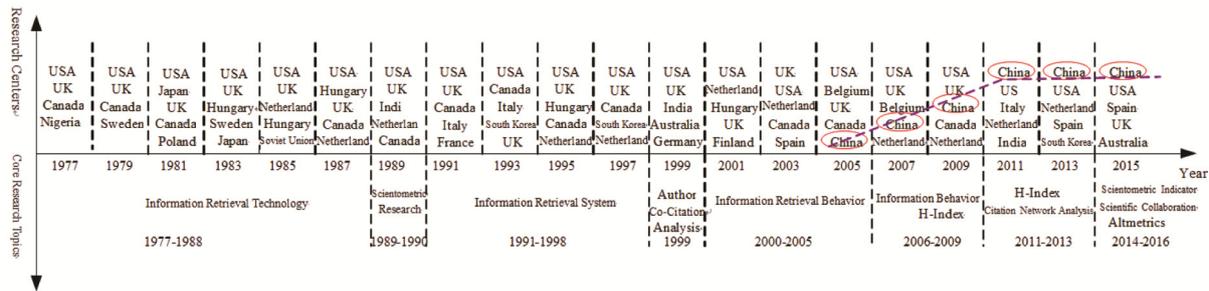


Figure 5. Evolution of the geographical distribution and topics of research centres.

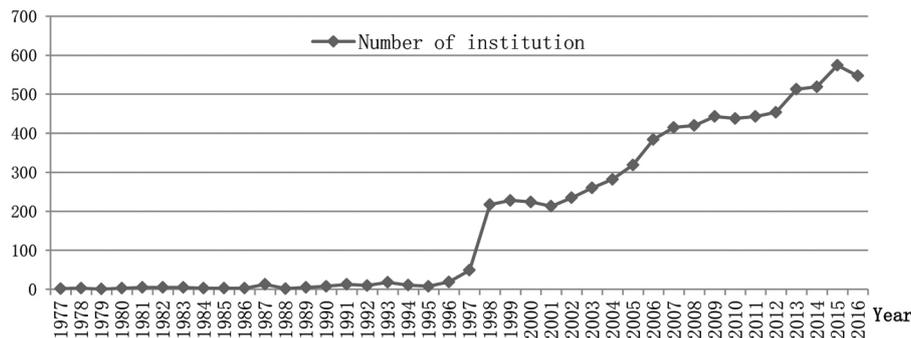


Figure 6. Variation in the number of institutions in IS domain.

output in the country as a whole. Hence, the country is listed as a central research country. In this case, the high-output institution is also likely listed as a central research institution. (2) A country has developed many institutions engaging in related research, but the individual publication output of each institution is relatively low, resulting in a high overall publication output in the country as a whole. Hence, the country is listed as a central research country. In this case, although the country is listed as a central research country, the institution in the country is not listed as a central research institution. (3) A country has developed many central research institutions that have many publications, resulting in a high overall publication output in the country as a whole. Hence, the country is listed as a central research country. Its high-output institution is also listed as a central research institution.

As far as central research institutions are concerned (Table 3), institutions as research centres of IS change every year. However, these institutions were located mainly in the USA and UK before the 1990s and included the University of Sheffield, Drexel University, Indiana University, and the University of North Carolina. Since the 1990s, an increasing number of institutions from other countries have become research centres. For example, McGill University in Canada, the University of Copenhagen in Denmark, the University of Amsterdam in The Netherlands, the Hungarian Academy of Sciences, and the University of Tampere in Finland all became centres of research between 1991 and 2008. From 2009 to

the present, major institutions as centres of research, including the University of Amsterdam in The Netherlands, the University of Wolverhampton in Britain, and the Max Planck Society in Germany have been playing active roles. In 2012, the Chinese Academy of Science became a centre for research for the first time, followed by Dalian University of Technology in 2013.

Research centres of IS were located in the United States during the 1970s and then migrated gradually to other countries. Between 1977 and 1987, countries as centres of research mainly included the US, Britain, Canada, Hungary and Sweden, but this group was still dominated by the US and Britain. Asian countries began to be involved in 1981, including Japan, South Korea, India, and China, among which Japan, South Korea and India were only included several times intermittently without outstanding performances. The status of China in this research field became prominent after 2005.

Hungary was ranked from 1983 to 1987. The Netherlands was ranked between 1985 and 2000, but did not get a core position until 2001. During 1988–2000, the dominant countries in IS were US and UK. Other prominent research countries included Canada, the Netherlands, and Italy. During 2001–2010, the dominant countries in IS were still US and UK, But other prominent research countries included the Netherlands, China, and Canada. The Netherlands had increased its position in the field of IS.

From 2011 to 2016, the leading countries of major research institutions of IS migrated from the US and Britain

RESEARCH ARTICLES

to China and the US. At present, the leading countries as major research centres of IS mainly include China, US, Spain, The Netherlands and Britain. China has been playing an important role with its research centres of IS since 2005, and ranks first in the research of key topics in IS.

Relationship between the changes in central research topics and the migration of central research country

Changes in central research topics have a close relationship with the rise of the country as a centre of research, and a change in the central research topic may lead to a change in the ranking of the country as a centre of research.

Conversely, a change in the ranking of the country as a centre of research may lead to a change in the central research topic. For example, the central research topic during 1997–1998 is IRS. The top ranking research centres were University of North Carolina and University of Arizona. The central research countries were the United States, Britain, Canada, The Netherlands and Italy. The central research topic during 2001–2005 was IRB, and although the research institutions changed every year, the

research centres remained the United States, Britain, The Netherlands, China and Canada. The central research topics during 2006–2009 were Information Behaviour and H-Index. No new countries had entered the ranks of centres of research, but the ranking of central countries had changed, for example, China's ranking had risen year by year. The central research topics have changed twice during the years 2011–2015, but the leading countries as centres of research have not changed. China remains first and the United States second. However, the ranking of countries as centres of research has changed. Since 2011, H-index has become a central research topic, and the country as a major centre of research has shifted to China and continues to the present.

Further looking into the formation characteristics of central research countries, we analyse the distribution of hot regions in the IS field through Google Fusion Tables. The number of research institutions extracted between 1997 and 2016 increased with each passing year (Figure 6).

We can visualize the distribution pattern of relevant research institutions from 1997 to 2016 as shown in Figure 7a. The hot regions in the IS field gradually shift

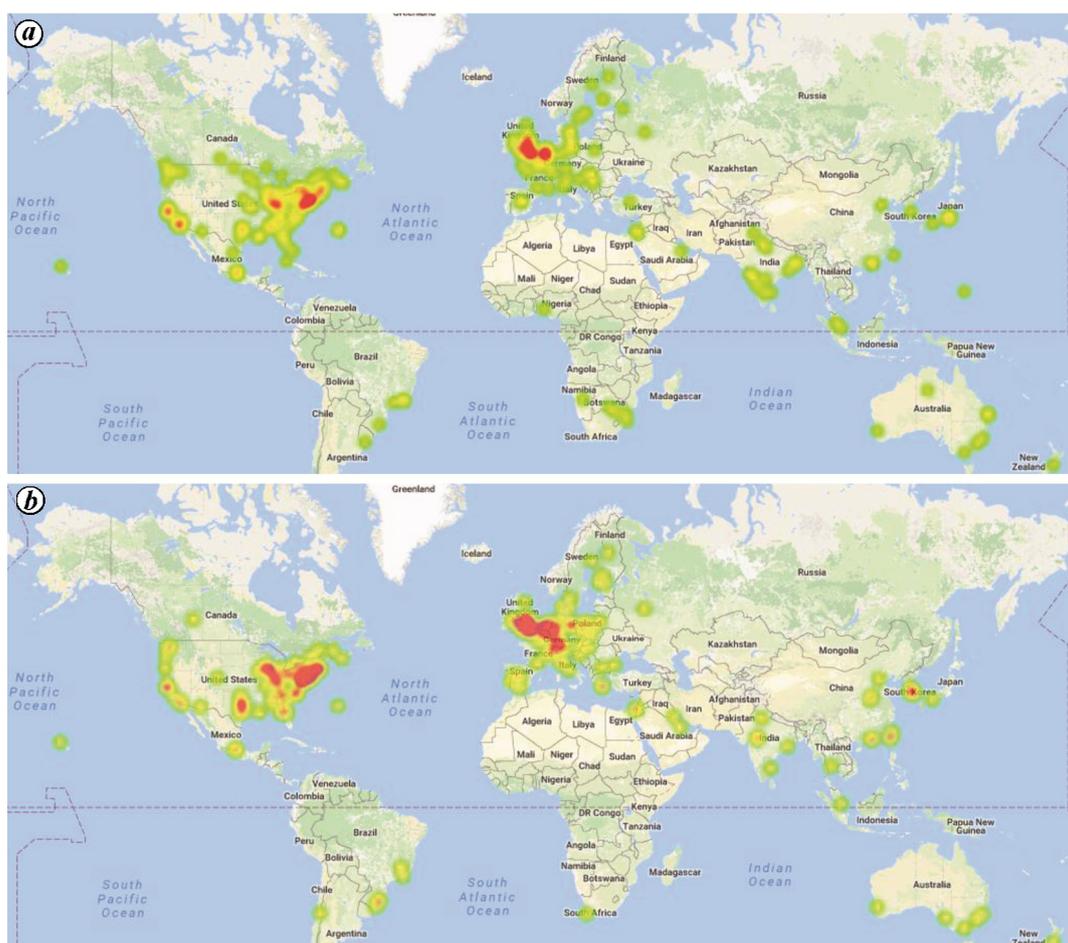


Figure 7. (Contd)

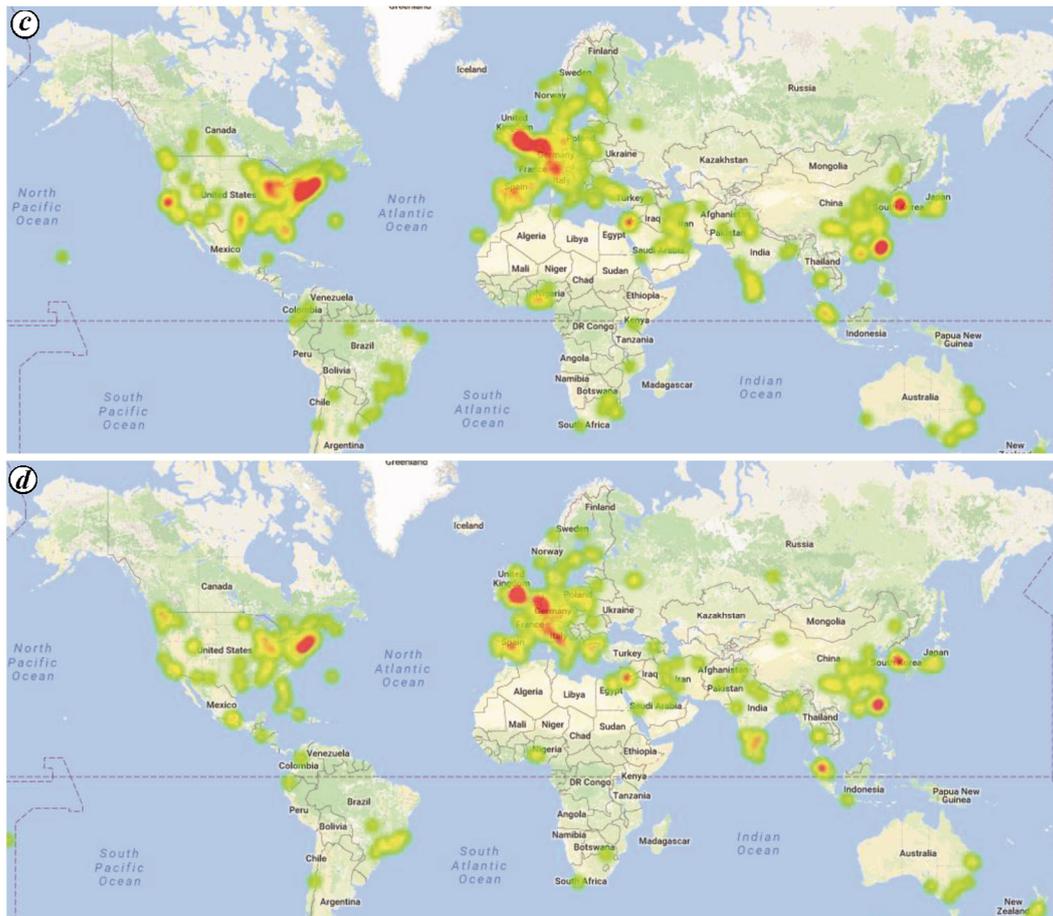


Figure 7. Distribution maps of institutions of information science. *a*, 1997; *b*, 2001; *c*, 2013; *d*, 2015.

from America and Europe to Asia, and three current hot regions have appeared (Figure 7*d*). The number of IS research institutions is growing year by year and is spreading geographically. To begin, the number of research institutions in America and Europe, such as the US and Britain increased and then spread first to neighbouring countries and then to Asian countries. Between 1997 and 2001, America and Europe were hot regions with many research institutions. They were mainly located in US, Britain, France and Germany. The leading institutions included Drexel University, the University of Chicago, the University of Illinois, the Hungarian Academy of Sciences and the University of London (Figure 7*b*).

In 2002, the hot regions began to shift to Asia, especially China and South Korea, which became research centres in 2005 and 2013 respectively. However, the role of South Korea is not that significant among the major centres of research. In 2009, the number of institutions conducting information science research in China ranked first in Asia, exceeded that of Britain, and reduced the gap between China and the US. Meanwhile, the hot regions in Europe started to shift from Britain, France, and Germany to Britain, The Netherlands, Belgium, and Spain in 2002 (Figure 7*c*). These countries all became

top-ranked research centres a few years later. Therefore, it can be concluded that all countries that are currently centres of research were hot regions in the past.

Relationship between the country as a centre of research and its academic influence

According to the distribution of institutions and countries, China, US, Britain, and The Netherlands were always ranked as centres of research between 2001 and 2006. To determine the relationships of the rank of a country as a centre of research with the total output of literature and the overall influence of the country, this study compared centres of research including US, China, and Britain with Japan and Sweden, which are not centres of research from three aspects: total output of literature, citation frequency and H-index. It should be noted that the measurement of these three indicators are all from the same sample data, that is, data of 12 core journals. The total output of literature from China, US and Britain exceeds 50% of the total output of literature in the world for every year between 2005 and 2014. The output of literature from the US and Britain decreased between 2005

and 2011, but rebounded and rose slightly after 2012. The output of literature from China rose during 2005–2014, exceeded that of Britain in 2009 and began to narrow the gap between China and US. The output of literature from China surpassed that of the US in 2014 for the first time. The output of literature from Japan and Sweden is low in general with a relatively stable number of publications (Figure 8). It can be seen that the countries that are centres of research have a high output of literature.

As for the total citation frequency of literature, the total number of citations of countries that are research centres all rise at first and then decline. The gap in the total number of citations of different centres of research is narrowing, indicating that the influence of literature of Chinese scholars has greatly increased. There is still a significant gap between the total number of citations of countries that are not centres of research and those of countries that are. The total number of citations of Japan rose first and then declined, while those of Sweden showed a declining trend on the whole (Figure 9).

The variation trend in the H-index is consistent with that of the total number of citations, but is different from that of the total output of literature. As time passes, the H-indices of different countries decrease. However, the H-indices of China, US, and Britain are much larger than those of Japan and Sweden, which are not centres of research. The H-index of China exceeded that of Britain

in 2009, and the gap between the H-index of China and that of the US also greatly decreased. The H-index of Britain rebounded in 2013 and caught up with that of China and US in 2014. It can be seen that China is playing an increasingly important role in IS, with the gap between China and the US as well as Britain in this field narrowing. China has become a core country in the field of IS (Figure 10). Countries that are centres of research enjoy great influence and achieve an H-index higher than 10 in 3 to 5 years after their scientific literature in the IS field is published.

Conclusions and discussions

This study employs information visualization tools including CiteSpace and Google Fusion Tables to explore the migration of centres of research in the IS field through scientometric analyses from several perspectives such as the evolution of research topics, and the distribution of leading research institutions and countries. We conduct an in-depth study on the relationship between central research topics and the countries behind them. Therefore, we arrive at the following conclusions:

(1) Polycentric tendency of the central research topics. The research topics from 1997 to 2006 in the IS field were all centred on information retrieval (techniques, systems and behaviour). However, this phenomenon changed after 2006. The central research topics began forming as 2 to 3, showing a tendency of polycentric development. At the same time, the period of prosperity of IS research topics was 12 to 8 years between 1980 and 1990, 6 to 4 years between 2000 and 2010, and 3 years between 2011 and 2016. This indicates that the pace of innovation in the IS field and the evolution of disciplines are accelerating.

(2) The central research topic is closely related to the migration of the country as centre of research. On one hand, the proliferation of the central research topic between countries caused an increase in the number of institutions researching this topic. This then forms the ‘hot

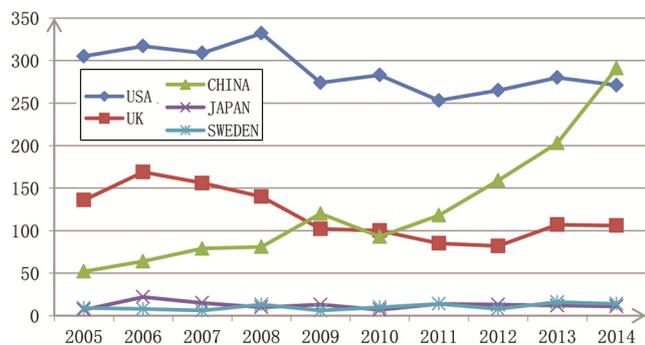


Figure 8. Variation trend of the number of published papers of five nations each year (2005–2014).

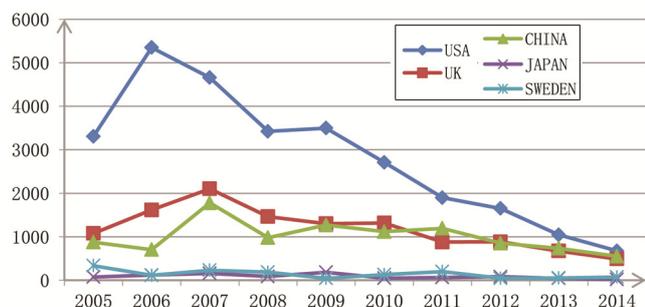


Figure 9. Variation trend of the citation frequency of documents published in five nations each year (2005–2014).

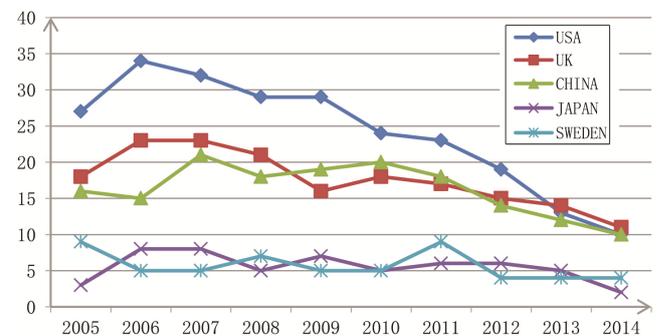


Figure 10. Variation trend of the H-index of five nations each year (2005–2014).

region', which can promote the formation of a country as centre of research, and change its ranking. On the other hand, the rise of central research countries can promote the increase of research institutions (volume of literature), breed new research topics and cause the shift of central research topics.

(3) A country as centre of research has a positive relationship with its academic influence. By analysing the conditions of central research countries through three indicators (output of literature, citation frequency and H-index) and conducting a comparative analysis of typical countries that are not research centres, this study shows that the central research country not only has high output of literature, but also has a high total citation frequency and H-index. These nations as research centres enjoy a high output of literature and also have great influence.

(4) China has become a research centre of IS. Research in the IS field started to shift to Asian countries rapidly after 2000. China became one of the top five IS research centres in 2005 for the first time, with its ranking rising year after year. As for research institutions, the Chinese Academy of Sciences and Dalian University of Technology became research centres in 2012 and in 2013 respectively.

The results have several implications for researchers and practitioners. Going by the theoretical implications, this study presents a visual domain analysis of the research centres of IS. Through information visualization, it reveals that the central IS research topics turn to be polycentric when development speeds up. Furthermore, this study reveals the characteristics of the transfer and the mechanism of the forming of IS research centres. For practitioners, this study provides a reliable historiography survey of IS research. The enhanced science mapping procedure introduced in this paper is applicable to the analysis of other domains of interest. Researchers can conduct visualized analyses of various specialties according to their needs, and effectively learn the dynamic evolution of these specialties.

However, there are some deficiencies in choosing these 12 journals. First, the criteria for choosing these journals was put forward by White and McCain according to the impact factor SSCI of JCR in 1993. Therefore, these journals cannot represent the current IS domain perfectly. Secondly, some of these 12 journals have ceased publication (without data), while the *Journal of Informetrics*, which launched in 2007, has developed into the core journal of the IS field at present. Therefore, the best solution is to choose different journals in different periods to define IS and its evolution more precisely. The two-step method can also be considered, that is, select the most representative core journals first, then conduct the journal co-citation analysis for the selected journals, and select journals as data samples according to the ranking of co-citation frequency of co-citation results. In this study, we do a visualization analysis of the transfer of research

centre in information science, mainly through the method of document co-citation analysis (DCA). Hence the analysis of data is the references of sample data, which is highly cited literature for each year. As time changes, high impact literature at different periods is distributed in different high influence journals. The literature also includes highly influential journals such as the *Journal of Informetrics*. Although journals such as *Journal of Informetrics* were not selected in the data selection of this paper, this did not make a significant difference in the results. CiteSpace can be used to label clusters automatically and avoid subjectivity in research topic analysis to a great extent. However, subjectivity cannot be completely avoided in choosing central research topics and defining the period of research prosperity.

Future research is expected to improve the method used in selecting literature data to define IS. For example, JCA can be used to select data with objective indicators such as the average impact factors of JCR in five years. In addition, for accurate and objective results, future research should try to minimize subjectivity when choosing core research topics. To achieve this, expert opinions can be taken into consideration.

1. Thelwall, M. and Maflahi, N., How important is computing technology for library and information science research? *Lib. Inform. Sci. Res.*, 2015, **37**(1), 40–50.
2. Tsay, M. Y., Knowledge flow out of the domain of information science: a bibliometric and citation analysis study. *Scientometrics*, 2015, **102**(1), 487–502.
3. Chu, H., Research methods in library and information science: a content analysis. *Lib. Inform. Sci. Res.*, 2015, **37**(1), 36–41.
4. Chen, C. M., Ibekwe-SanJuan, F. and Hou, J. H., The structure and dynamics of cocitation clusters: a multiple-perspective co-citation analysis. *J. Am. Soc. Inf. Sci. Technol.*, 2010, **61**(7), 1386–1409.
5. Yuasa, M., Center of scientific activity: its shift from the 16th to the 20th century. *Jpn. Stud. Hist. Sci.*, 1962, **1**(1), 57–75.
6. Zhao, H. D., *Introduction to Scientific Competence*, Science Press, Beijing, 1984.
7. Zhao, H. Z. and Jiang, G. H., Shifting of world's scientific centre and scientists social ages. *Scientometrics*, 1985, **8**(1–2), 59–80.
8. Liang, L. M., Feng, Y. and Wu, Y. S., Shifts in the world science centre: space-time characteristics and disciplinary analysis. *Interdiscip. Sci. Rev.*, 2000, **25**(3), 227–232.
9. Levitt, J. M. and Thelwall, M., Long term productivity and collaboration in information science. *Scientometrics*, 2016, **108**(3), 1103–1117.
10. Zhao, D. Z. and Strotmann, A., The knowledge base and research front of information science 2006–2010: an author cocitation and bibliographic coupling analysis. *J. Assoc. Inform. Sci. Technol.*, 2014, **65**(5), 995–1006.
11. White, H. D. and McCain, K. W., Visualizing a discipline: an author co-citation analysis of information science 1972–1995. *J. Am. Soc. Inf. Sci.*, 1998, **49**(4), 327–355.
12. Astrom, F., Changes in the LIS research front: time-sliced co-citation analyses of LIS journal articles, 1990–2004. *J. Am. Soc. Inf. Sci. Technol.*, 2007, **58**(7), 947–957.
13. White, H. D., Pathfinder networks and author cocitation analysis: A remapping of paradigmatic information scientists. *J. Am. Soc. Inf. Sci. Technol.*, 2003, **54**(5), 423–434.

14. Zhao, D. Z. and Strotmann, A., Evolution of research activities and intellectual influences in information science 1996–2005: introducing author bibliographic-coupling analysis. *J. Am. Soc. Inf. Sci. Technol.*, 2008, **59**(13), 2070–2086.
15. Zhao, D. Z. and Strotmann, A., Information science during the first decade of the web: an enriched author co-citation analysis. *J. Am. Soc. Inf. Sci. Technol.*, 2008, **59**(6), 916–937.
16. Zhao, D. Z. and Strotmann, A., Comparing all-author and first-author co-citation analyses of information science. *J. Inform.*, 2008, **2**(3), 229–239.
17. Ibekwe-SanJuan, F., Information Science in the web era: a term-based approach to domain mapping, In *Annual Meeting of American Society for Information Science and Technology*, Vancouver, Canada, 6–11 November 2009, pp. 1–13.
18. Klavans, R., Persson, O. and Boyack, K. W., Coco at the copacabana: introducing co-cited author pair co-citation (coco) analysis. In *Proceedings of the International Conference on Scientometrics and Informetrics*, Rio de Janeiro, Brazil, 2009.
19. Yang, S., Han, R., Wolfram, D. and Zhao, Y., Visualizing the intellectual structure of information science (2006–2015): introducing author keyword coupling analysis. *J. Informetr.*, 2016, **10**(1), 132–150.
20. White, H. D. and Griffith, B. C., Author co-citation: a literature measure of intellectual structure. *J. Am. Soc. Inf. Sci.*, 1981, **32**(3), 163–171.
21. Janssens, F. *et al.*, Towards mapping library and information science. *Inform. Proc. Manage.*, 2006, **42**(6), 1614–1642.
22. Janssens, F. *et al.*, A hybrid mapping of information science. *Scientometrics*, 2008, **75**(3), 607–631.
23. Besselaar, P. V. D. and Heimeriks, G., Mapping research topics using word-reference co-occurrences: a method and an exploratory case study. *Scientometrics*, 2006, **68**(3), 377–393.
24. Milojević, S., Sugimoto, C. R., Yan, E. and Ding, Y., The cognitive structure of library and information science: analysis of article title words. *J. Am. Soc. Inf. Sci. Technol.*, 2011, **62**(10), 1933–1953.
25. Jeong, Y. K., Song, M. and Ding, Y., Content-based author co-citation analysis. *J. Informetr.*, 2014, **8**(1), 197–211.
26. Chang, Y. W., Huang, M. H. and Lin, C. W., Evolution of research subjects in library and information science based on keyword, bibliographical coupling, and co-citation analyses. *Scientometrics*, 2015, **105**(3), 2071–2087.
27. Lei, L. and Yan, S., Readability and citations in information science: evidence from abstracts and articles of four journals (2003–2012). *Scientometrics*, 2016, **108**(3), 1155–1169.
28. Khasseh, A. A., Soheili, F., Moghaddam, H. S. and Chelak, A. M., Intellectual structure of knowledge in iMetrics: a co-word analysis. *Inform. Proc. Manage.*, 2017, **53**(3), 705–720.
29. Yang, S. and Wang, F., Visualizing information science: author direct citation analysis in China and around the world. *J. Inform.*, 2015, **9**(1), 208–225.
30. Chen, C. M., CiteSpace II: detecting and visualizing emerging trends and transient patterns in scientific literature. *J. Am. Soc. Inf. Sci. Technol.*, 2006, **57**(3), 359–377.
31. Chen, C. M., Science mapping: a systematic review of the literature. *J. Data Inform. Sci.*, 2017, **2**(2), 1–40.
32. Bradley, E. S., Roberts, D. A. and Dennison, P. E., Google Earth and Google Fusion Tables in support of time-critical collaboration: mapping the deepwater horizon oil spill with the AVIRIS airborne spectrometer. *Earth Sci. Inform.*, 2011, **4**(4), 169–179.
33. Signore, A., Mapping and sharing agro-biodiversity using Open Data Kit and Google Fusion Tables. *Comput. Electron. Agric.*, 2016, **127**, 87–91.
34. Green, L. Y., Mikhailova, E. A., Post, C. J., Darnault, C. C. J. G., Bridges, W. C. and Schlautman, M. A., A cloud-based spatial-temporal inventory for sustainable urban soil management. *Urban Ecosyst.*, 2016, **19**(2), 811–822.
35. Mingers, J. and Leydesdorff, L., A review of theory and practice in scientometrics. *Eur. J. Oper. Res.*, 2015, **246**(1), 1–19.
36. Bornmann, L. and Haunschild, R., How to normalize Twitter counts? A first attempt based on journals in the Twitter Index. *Scientometrics*, 2016, **107**(3), 1405–1422.
37. Saif, H., He, Y. L., Fernandez, M. and Alani, H., Contextual semantics for sentiment analysis of Twitter. *Inf. Process. Manage.*, 2016, **52**(1), 5–19.
38. Ozdikis, O., Ogurtuzun, H. and Karagoz, P., Evidential estimation of event locations in microblogs using the Dempster-Shafer theory. *Inf. Process. Manage.*, 2016, **52**(6), 1227–1246.
39. Haustein, S., Bowman, T. D., Holmberg, K., Tsou, A., Sugimoto, C. R. and Lariviere, V., Tweets as impact indicators: examining the implications of automated ‘bot’ accounts on Twitter. *J. Assoc. Inf. Sci. Technol.*, 2016, **67**(1), 232–238.
40. Laylavi, F., Rajabifard, A. and Kalantari, M., Event relatedness assessment of Twitter messages for emergency response. *Inform. Proc. Manage.*, 2017, **53**(1), 266–280.
41. Klavans, R. and Boyack, K. W., Using global mapping to create more accurate document-level maps of research fields. *J. Am. Soc. Inf. Sci. Technol.*, 2011, **62**(1), 1–18.
42. Hirsch, J. E., An index to quantify an individual’s scientific research output that takes into account the effect of multiple co-authorship. *Scientometrics*, 2005, **85**(3): 741–754.

ACKNOWLEDGEMENT. This research was supported by the National Social Science Foundation of China under Grant 17BGL031.

Received 15 August 2018; revised accepted 17 October 2018

doi: 10.18520/cs/v116/i3/422-436