Collaboration - Changing the Global Landscape of Science

Proceedings of
10th International Conference on
Webometrics, Informetrics and Scientometrics &
15th COLLNET Meeting 2014

September 3-5, 2014
Technische Universität Ilmenau, Germany

Edited by
Bernd Markscheffel • Daniel Fischer •
Daniela Büttner • Hildrun Kretschmer

THOMSON REUTERS

TECHNISCHE UNIVERSITÄT
ILMENAU

Elsevier
Research Intelligence

Collaboration – Changing the Global Landscape of Science

Proceedings of
10th International Conference on
Webometrics, Informetrics and Scientometrics &
15th COLLNET Meeting 2014

September 3-5, 2014
Technische Universität Ilmenau, Germany

Edited by

Bernd Markscheffel,

Daniel Fischer,

Daniela Büttner and

Hildrun Kretschmer

Collaboration – Changing the Global Landscape of Science

Proceedings of 10th International Conference on Webometrics, Informetrics and Scientometrics & 15th COLLNET Meeting 2014

Bernd Markscheffel,
Daniel Fischer,
Daniela Büttner and
Hildrun Kretschmer

Technische Universität Ilmenau
Fakultät für Wirtschaftswissenschaften und Medien
Institut für Wirtschaftsinformatik
P.O. Box 100565
98684 Ilmenau
Germany

bernd.markscheffel@tu-ilmenau.de
daniel.fischer@tu-ilmenau.de
daniela.buettner@tu-ilmenau.de
kretschmer.h@onlinehome.de

URN: urn:nbn:de:gbv:ilm1-2014200150

Ilmenau, 2014

Index

Invited Papers

Journal Impact Factor Reflects Citedness of the Majority of the Journal Papers

Eugene Garfield* and Alexander Pudovkin**

*ThomsonReuters Scientific, 1500 Spring Garden Street, Philadelphia, PA 19130-4067
eugene.garfield@thomsonreuters.com

**Institute of Marine Biology, Vladivostok 690041, Russia
aipud@mail.ru

Introduction

The literature on Journal Impact Factors (JIF) is quite rich. A special issue of "Scientometrics" is dedicated to the question (Braun, 2012). Recent literature is discussed by Brody (2013). Though the bulk of literature considers drawbacks and shortages of the JIF (see the issue of "Scientometrics"), there seems to be no better quantitative journal characteristic (Brody, 2013). It still is very popular among publishing scientists (Crotty, 2013). The main suggested drawbacks of JIF are 1) their presumptive dependence from only a few highly cited papers published in the journal, and 2) possibility to manipulate their values by a) taking account of all the citations to the journal, while disregarding the uncitable items (editorial materials, corrections, letters, notes, etc.), which actually may be cited and sometimes are cited; we call this "numerator/denominator trick"; b) encouraging or even requiring authors to cite papers of the journals. All these may inflate IF value of a journal, disproportionally to the actual citedness of the majority of the journal's papers.

To test the idea that JIF does reflect the citedness of the majority of journal's papers (rather than depends on only a few highly cited ones) we calculated coefficients of correlation between the JIF and the citation score of the median (by citation score) paper of the journal for journals of 5 JCR specialty categories. Table1 gives a schematic explanation of why the median rather than the arithmetic mean is informative in this matter.

Table 1. A schematic numerical example showing differences of the arithmetic mean and the median: independence of the median on the maximal values in the set.

Citation Rank of papers	Journal 1	Journal 2
1	1000	24
2	500	22
3	20	20
4	18	18
5	16	16
6	14	14
7	12	12
8	10	10
9	8	8
10	6	6
mean	160.4	15.0
median	15	15

The table shows ranked distributions of citation numbers of papers in two journals. In the 1st journal the distribution if very skewed: a couple of high values followed by a tail of much lower values. In the 2nd journal the distribution is symmetrical. Though there is a great difference between the means of the two distributions, the medians are the same, being independent of the high values in the top of the distribution. It is well known that the distributions of citation numbers are skewed (Seglen, 1992). The JIF is an arithmetic mean, so it seems it should depend on only a few highly cited papers. But this is not shown empirically. The median paper is in the middle of the ranked (by citation numbers) list of papers. It is quite far from the top cited papers, all the papers above the median have more citations. If JIF is correlated with the citation number of the median paper, it means that it is correlated with the citation numbers of half of journal's papers. Thus, if the correlation is significant, it means that at least half of journal papers significantly contribute to the IF value.

Data Collection

JIF values from JCR, 2012 were imported into MS-Excel table. Then, for each journal included in the category we obtained the citation scores of the median paper. We limited the documents considered to “articles”, “reviews”, and “proceeding papers”. This was done on July 15-17, 2013. For that we searched in the WoS for all the papers (note the above mentioned limitation) of the journal published in 2010-2011, sorted them by “times cited” (during 2010 to July 15-17, 2013) and looked for the median paper. The obtained values we entered into the same MS-Excel table. Thus, we had IF values for all the journals of the category and the cite numbers of the median paper. The latter were usually larger than the IF value. This is due to the differences in the time periods for the two values: IF, 2012 is the average citation scores of papers of 2010-2011 accumulated in one year, 2012, while the median citation scores we got are accumulated during 3.5 years: January 1, 2010 to July, 15, 2013.

Results

The summary of the data obtained are given in Table 2. One can see that coefficients of correlation, r are very high, close to 1. This means that IF of a journal reflects the overall citedness of the journal, the citedness of the majority of its papers. If IF value of a journal would have been caused by only a few highly cited papers, specifically solicited by the editors of the journal or just happened to occur in the journal there would be no correlation with the citation score of the median paper which may be quite far below from the top cited ones.

Table 2. Summary of the data obtained

JCR Category	r	Number of Journals in the Category	Actual* Number of Journals in the Category	Median Number of Papers in the Journal
Physics, Condensed Matter	0.994	68	66	483
Genetics & Heredity	0.990	161	159	157
Marine & Freshwater Biology	0.976	100	97	122.5
Multidisciplinary Sciences	0.997	56	55	142
Information Science & Library Science	0.879	84	83	62.5

*For some journals data in JCR or WoS were incomplete and these journals were omitted.

It should be stressed, that for all the 5 specialty categories we observe strictly linear relationship between IF values and numbers of cites for the median paper. Fig.1 shows scatter diagram for 66 journals of the category “Physics, Condensed Matter”.

Certainly, there may be and actually are such cases, when high IF of the journal is due to the occurrence of a very few highly papers. But we located only one such journal out of about 500 considered. Fig.2 gives the scatter diagram of 30 top IF journals. One can see that there is an outlier, the journal CA-A CANCER JOURNAL FOR CLINICIANS (to the right of the diagram). This journal is the top IF journal, its IF = 153.459, though there are only 36 cites to its median paper, which is only the 15th from the top. Its high IF is due to 2 extremely highly cited papers giving world statistics for the occurrence of cancer. Citation numbers for the 5 top cited papers in this journal (in 2010-2011) are 4531, 3098, 1030, 208, 171. If one computes correlation "r" for the 30 journals (including CA-Cancer J Clin) one obtains r = 0.029, which is insignificant. Though, if one omits this evident outlier, the r = 0.704, which is significant. Interestingly, this outlying position of CA-A CANCER JOURNAL FOR CLINICIANS is seen in the "bubble diagram" generated by Davis (2013) for 10-year WoS data set, 2003-2012.

Discussion

Our finding of very high correlation of IF values and the median citation rates does provide convincing evidence that IF values are not due to a few highly cited papers but rather characterizes the majority of the journal's papers. It is not a predictor of the future citation rate of a paper published in the journal: the citation score for the paper might happen to be much lower than the median. The data show that the IF is a predictor of the citation score for an AVERAGE (median) paper of the journal (and of all the papers above the median).

If editors of many journals did use the "numerator/denominator" trick, then there would be many outliers like that in Fig.2 (CA-Cancer J Clin) that is the journals with high IF and relatively low median citation number. Though, we observed this case only for one journals among about 500 considered.

Some authors claim that editors encourage or even require authors to cite the journals thus artificially inflating journal's IF. We looked into the matter calculating correlation coefficient between journals' IF and its self-citedness, which is provided by JCR ("Journals Self Cites" button, percent of "Self Cites to Years Used in Impact Factor Calculation"). We did it for two specialty categories: "Physics, Condensed Matter", and "Genetics & Heredity". We got significant, but negative "r": -0.335 and -0.256. Thus, excess of self-citation is related with low IF rather than high. Possibly, some editors do use this trick to enhance the IF value, but the majority of journals with excessive self-citation have IF less than 2.5 (see Fig.3 and Fig.4). Some of these journals are national (thus high self-citedness is quite understandable), some are "narrow specialty" monopolists, hence high self-citation. Besides, JCR provides JIF without self-citation. So, if there are some doubts about this way of inflation, one may use JIF value, corrected for self-citation.

To reiterate, It would be advisable not to use IF values as proxies for citation counts of individual papers. The IF, however, is an indication of the standing of the journal, its prestige or authority. Even in high IF journals there are some poorly cited papers. But even these poorly cited papers are usually good, professional papers. They went through the sieve of thorough refereeing and editing. It is this "refereeing sieve" that justifies using JIF values in evaluation procedures, especially for recently published papers. Thus, it seems quite reasonable to use JIF for rating recently published papers which have not yet accumulated due cites. It should be taken into account that JIF differ greatly among specialties. To make comparisons more fair one may use rank-normalized IF within the specialty category as suggested by Pudovkin and Garfield (2004). It seems especially adequate procedure for mass monitoring and weeding out weak position candidates or grant applicants, whose publications appeared in obscure journals having extremely low IF. Of course, JIF statistics should not be a single characteristic for judgment.

Literature

Braun, T. (Ed.) (2012). Special discussion issue on journal impact factors [Special issue]. *Scientometrics*, 92(2).

Brody S. 2013. Impact factor: Imperfect but not yet replaceable. *Scientometrics*, 96, p. 255-257.

Crotty D. *The Persistent Lure of the Impact Factor–Even for PLOS ONE* http://scholarlykitchen.sspnet.org/2013/07/30/the-persistent-lure-of-the-impact-factor-even-for-plos-one

Davis P. 2013. *Dynamic Visualization of Biomedical Journals 2003-2012*. http://wp.me/pcvbl-8Fm

Pudovkin Al, Garfield E, "Rank-Normalized Impact Factor: A way to compare journal performance across subject categories". *Proceedings of the 67th Annual Meeting of the American Society for Information Science & Technology*, vol:41, p.507-515, 2004.

Figures

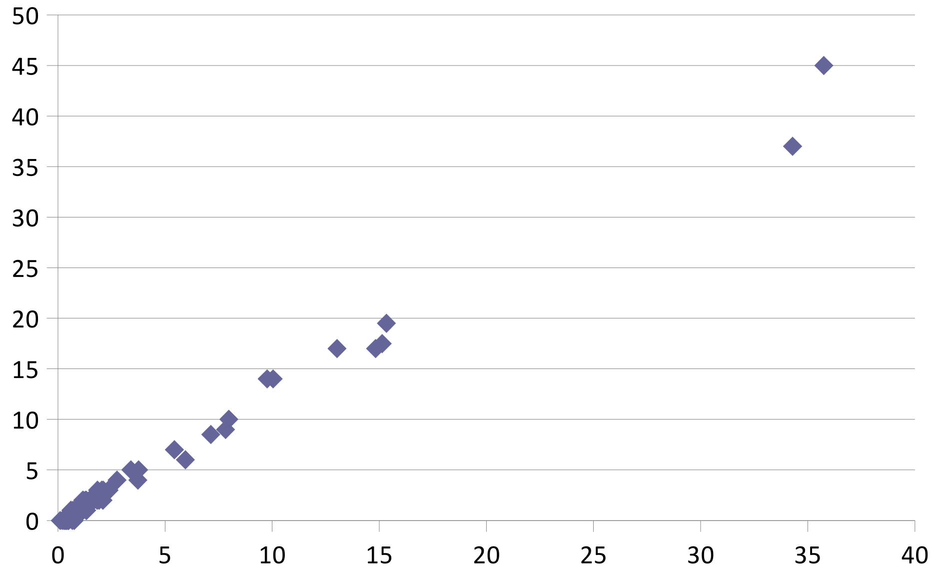


Fig.1. Scatter diagram showing relationships between the values JIF, 2012 (on the abscissa) and citation score of a median paper (on the ordinate), accumulated during the period of January, 2010 to July 15, 2013 for the 66 journals in the JCR specialty category "Physics, Condensed Matter".

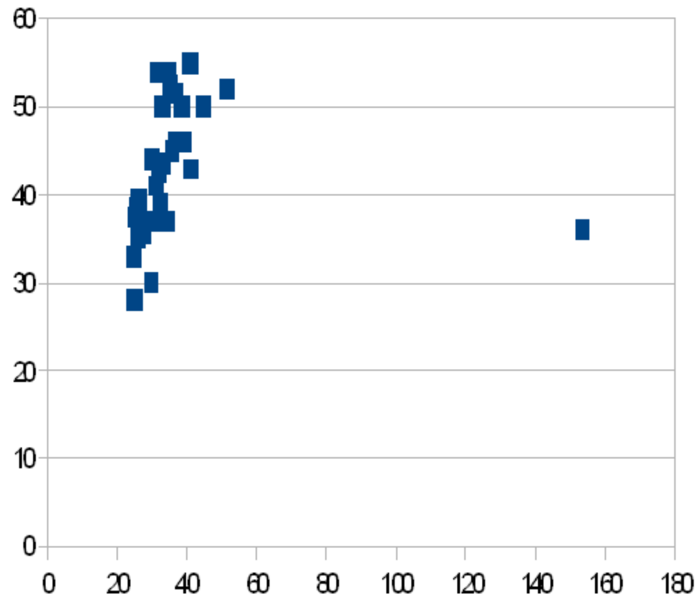


Fig.2. Scatter diagram showing correlation between IF values (on the ordinate) of 30 top IF journals and the citation score for a median paper. The details are the same as in Fig.1.

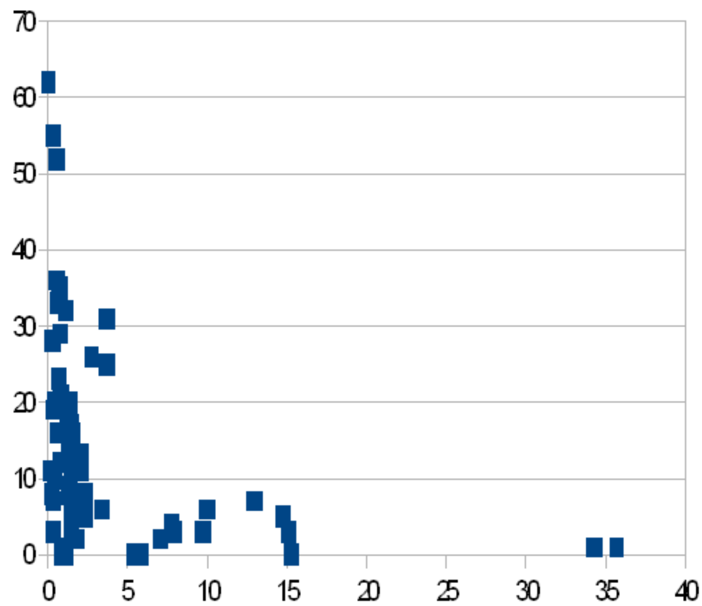


Fig.3. Correlation of Journal Impact Factor (on the abscissa) and self-citedness of the journal (per cent of self cites, on the ordinate) for journals of JCR category "Physics, Condensed Matter".

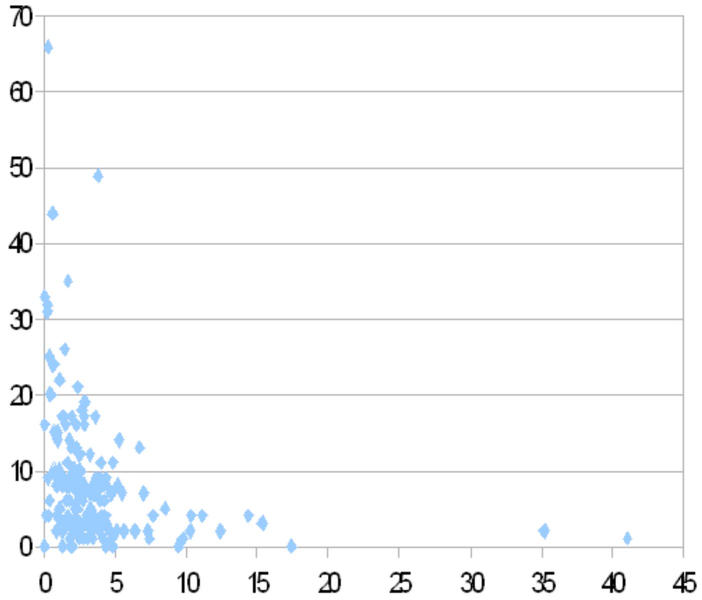


Fig.4. Correlation of Journal Impact Factor (on the abscissa) and self-citedness of the journal (per cent of self cites, on the ordinate) for journals of JCR category "Genetics & Heredity"

A Scientometric Study on Collaboration between Academia and Industry – Case studies of Chinese leading universities and companies

Weiping Yue

Thomson Reuters

Since the beginning of 21st century, China has promoted indigenous innovation into China's national strategy. Domestic companies are considered as a key driver for indigenous innovation. However, innovative R&D capabilities of Chinese local companies vary and still need to be enhanced. To accelerate innovation outcomes and improve business performance, many domestic companies have adopted open innovation and actively collaborated with universities and research institutes to leverage their research resources and capacities. On the other hand, universities also seek for research funding from industry and commercialize their academic research outputs. This research aims to investigate the collaboration between academia and industry in China at institutional level and to identify influential factors underlying the collaboration.

In this study, collaboration between academia and industry covers various forms of engagement, i.e. joint research, contract research, patent transfer and technology transfer. Indicators reflecting such collaboration include number of co-authored papers, number of co-owned patents, number of research papers funded by industry, university research funding from industry, number of contracts and total incomes in terms of patent transfer at universities, and number of contracts and total incomes in terms of technology transfer. Data was collected from Web of Science core collection, Derwent World Patent Index, InCites, and Compilation of Statistics on University S&T Resources provided by Ministry of Education of China. A scientometric analysis was applied to data collected from leading universities and companies in China, who are ranked as top entities in terms of total number of inventions in the white paper of Research & innovation performance of the G20.

Preliminary results showed that more than 70% of research papers authored by Huawei, ZTE and SINOPEC are joint research outputs with academia. Over 40% of collaborative papers published from 2000 to 2013 are from the most recent three years, which demonstrated an increasing trend of university and academia collaboration. In terms of published papers funded by each of the three companies, the majorities of the output are also from the university research. SINOPEC has published about 1460 papers with universities since year 2000 while Huawei's collaboration with academia has reached to 28 countries and territories. But in the analysis of patents, it is found that the percentage of co-owned patents between these three companies and universities are quite low. SINOPEC's co-owned patents have the biggest share within the three companies, but it is only 3.2% of its total applied patents from 2000 to 2013.

From the university perspective, the co-authored papers with industry for Tsinghua University have also increased steadily in the last ten years. However, the percentage of industry collaboration papers is only a small portion of total number of university research papers in each year. A search of funding acknowledgement shows even smaller portion of papers funded by industry.

A closer look at the patent portfolio of Tsing Hua University revealed a substantial fraction of inventions that are co-assigned and many different companies are involved. The most prominent companies are Hon Hai, Tongfang NucTec, Beijing Visionox, and Capital Bio/Boao. Both Tongfang NucTec and Capital Bio/Boao are companies owned by Tsinghua University. Co-assignment with industry began around year 2000 and has increased steadily. Commercial entities involved with Tsinghua University drive filing outside China, which shows there is a possible market for the invention elsewhere. The interaction with companies drives external interest in the form of citations, and the quality scores for patents with commercial co-owners are generally a little higher.

Tsinghua University's research income from industry in terms of total amount of investment has increased rapidly from 2008 to 2012. But bearing in mind the increase of total research funding of the university, the share of research income from industry at Tsinghua University has remained at about 40% at Tsinghua University. The number of contracts, total incomes, and average income per contract of technology transfer at Tsinghua University ranked first in the C9 university group in 2012. Although the number of patents and the number of contracts of patent transfer at Tsinghua University were lower than another two C9 university in 2012, its total incomes and average income per patent transfer contract performed outstandingly within the C9 group.

In this study, it is not surprising to find that research papers authored by leading companies in China mainly come from the collaboration with academia. Providing research funding to universities has become a main form of the collaboration. However, because of the future commercial benefits, companies intend to own patents by themselves, but not to co-own the patents with their research partners at universities. It is critical for university research administrators to strive for the rights and interests when making any types of collaboration agreement with industry.

The case study of Tsinghua University shows that researchers have applied more patents with commercial entities than published papers together on scientific journals. This phenomenon is reflected by both the number of patents co-owned with industry partners and the level of involvement of companies. This might due to the fact that number of patents has become an indicator for national research assessment practice, and also that universities provide incentives for researchers to apply for patents and pay annuity.

It is found that for key technologies developed at Tsinghua University, the government and university have supported and invested to establish Tsinghua Holding Co. Ltd to further commercialize its research outcomes. Other top universities such as Peking University and Zhejiang University also have their holding group companies, which are often large scale. This might be special for China as universities overseas tend to encourage the creation of startup companies.

Industry and academia collaboration in China has increased steadily and tend to continually expand. But the level of collaboration has been mainly influenced by some institutional factors, for example, company business strategy and R&D strategy, university's research capability, university awarding system, and various financial incentives and policy support from university. In addition to the institutional factors, China central government still has played an important role in driving and accelerating innovation. In 2011, Ministry of Education and Ministry of Finance announced Higher Education Innovative Capacity Improvement Scheme (also called as project 2011) to accelerate the establishment of China as an innovative country generating high quality research outcomes, using collaborative partnerships as the key mechanism. Project 2011 was in light of former President Hu Jintao's speech at Tsinghua University in 2011, where he challenged Chinese universities to increase both their innovation capacity and the application of their research outcomes. Since then many collaborative innovation centers have been created and the government provided its first investment to selected centers in 2013. It will take several years to observe and assess their research outcomes and impact.

Full Papers

A Survey on Collaboration rate of authors in producing Scientific Papers in Quarterly Journal of Information Technology Management during 2009-2014

Amir Reza Asnafi* and Maryam Pakdaman Naeini**

*Faculty member of Shahid Beheshti University, Information Science and Knowledge Department, Velenjak, Tehran, I.R.Iran
a_asnafi@sbu.ac.ir

**PhD student of Information Science and Knowledge, Payame Noor Mashahd University, and librarian IIEES
m.pakdman@gmail.com

Abstract

Current research wants to determinate the collaboration among authors who published papers in Journal of Information Technology Management during 2009-2014 in Iran. Findings revealed that scholars had fewer trends to publish one author paper in Journal of Information Technology Management. 475 authors published 158 papers in journal of Information Technology Management during 2009-2014. Current research revealed that 6 papers were individual and 152 papers were group. In average for each paper, 3.01 authors had collaboration. Findings indicated that author's collaboration coefficient in Journal of Information Technology Management was 0.608 that means this is a desirable status. Current research revealed that papers of Journal of Information Technology Management that authors had trend to collaboration and group papers. Sharing in knowledge, resources and responsibilities are considered in most of scientific disciplines, so group works shape most of publication.

Keywords: Journal of Information Technology Management, Collaboration rate, Authors Collaboration Coefficient

Introduction and Research Questions

Journal of Information Technology Management is an Iranian journal that is published by Tehran University in Iran from 2009 and focuses on fields like: Knowledge management, Information Technology and related fields. The main objective of current research was determination of collaboration among authors who published papers in Journal of Information Technology Management during 2009-2014.

Current research has a glance on collaboration rate among authors of Journal Information Technology Management in Iran during 2009-2014 and wants to answer these questions:

1. How many authors partnered for publishing papers in Journal of Information Technology Management?
2. Which Universities had the most publications in Journal of Information Technology Management
3. How much is the Authors Collaboration Coefficient in studied journal?
4. Which papers of Journal of Information Technology Management in Islamic Science Citation (ISC) database?
5. Which papers of Journal of Information Technology Management are the highly cited articles in ISC?
6. Who are the most active authors in studied journal?

Literature Review

Noruzi and Alimohammadi (2006) measured the number of contributions by Iranian librarians and information professionals published in international journals indexed by the ISI citation indexes. It is concluded that the number of papers published by Iranian librarians and information professionals is low, although there is an increase since 1992. The study also shows that the scientific collaboration between Iranian information professionals and between them and their international peers is weak. Writing articles in English is recommended to increase the rate of contribution of Iranian LIS professionals in the international level.

Osareh and Wilson (2002) in a research undertaken to survey the rate of international collaboration in the scientific works of the Iranians in the area of science citation index during the years 1995-1999 and in comparison with their previous study on the same theme found that the scientific works of the Iranians in science and technology in three five year periods in this area has increased. Iran has increased its publications by two fold in the first two periods and by 2.8 times in the third period. The greater part of the Iranian's international collaboration in these three periods has been with American and British co-authors and collaboration with the authors of other nations has also had a significant increase. Osareh and Marefat (2005) in a research surveyed the growth and development of the articles submitted by Iranian researchers in foundation sciences and inter- medicinal areas to the medical science information network Medline in the years 1976 to 2003 and identified the Iranian universities, journals and researchers who had produced the most scientific articles and indicated those subject areas which these researchers were wore interested in. The results of this study indicated that articles and materials submitted to Medline by Iranian researchers had increased significantly so that during the period under research 2695 articles from 9373 coauthors has been published where the average number of authors collaborating on an article was 3.4 authors. In the international scale there have been many researches on collaboration in the production of scientific material. Sarrafzadeh(2000) in her masters thesis studied the state of the Iranian articles indexed on the CAB and Agris databases since the beginning till 1997 with the aim of the determination of the share of the Iranian articles from the total number of the articles that had appeared on these databases and the clarification of the extent of the collaboration of each of the nations educational and research centres in the production of the articles present in these databases. The results indicate a reduction in the number of Iranian articles submitted after the Islamic revolution in Iran (1979) which he attributes to the occurrence of events such as the Iranian Revolution, The closure of the universities and the Iran-Iraq war. However, from the nineties onward there has again been an increase in the appearance of Iranian articles on these databases. Other data indicated that from the 47 centres which had contributed more than 5 articles to these databases, The University of Tehran had the biggest share and The Semnan Agricultural Research Centre the least. Liang, Kretschmer, Guo, Beaver (2001) had a study on age structures of scientific collaboration in Chinese computer science. Analysis reveals some special age structures in scientific collaboration in Chinese computer science. Most collaborations are composed of scientists younger than thirty-six (Younger) or older than fifty (Elder). For two-dimensional collaboration formed by first and second authors, Younger-Elder and Younger-Younger are the predominant age structures. For three-dimensional collaboration formed by first, second and third authors, Younger-Younger- Elder and Younger-Younger-Younger are the most important age structures. Collaboration between two authors older than 38 amounts to only 6.4 percent of all two-person collaborations. Collaboration between two middle-aged scientists is seldom seen. They suggest a tentative explanation based on analyses of the age composition of all authors, the age distributions of the authors in different ranks, and the name-ordering of authors in articles written by professors and their students. Gupta & Dhawan (2007) reviewed the present status of Indian physics, particularly with regard to the nature of research system, nature of institutions involved, type of education available and outturn at postgraduate and Ph.D level, the extent of extra-mural funding support available from various agencies, and the nature of professional organizations involved Analyses the growth of Indian physics output, as reflected in mainstream international journals covered in Expanded Science Citation Index (Web of Science) during 1993-01. Discusses the various features of Indian physics research output, such as growth, institutional publication productivity, nature of collaboration, and the quality and impact of its research output.

Results of Hayati and Didegah paper (2010) showed that Iranian researchers have had scientific collaboration with 115 countries, and that their numbers have increased between 1998 and 2007. The results also showed that the number of domestic articles per year was 2-3.5 times more than international ones. Investigating international collaboration in different subject areas revealed that geosciences had the biggest number of publications co-authored internationally. Iran's main partners were the USA, Canada, and UK, respectively. European researchers were the main counterparts of Iranian researchers. In addition, Iranian researchers had mostly co-published with their colleagues in advanced countries. Among Iranian universities and research institutions, the University of Tehran had the highest collaboration at the international level. The results revealed that the average number of citations received by international co-authored publications was more than those received by domestic co-authored publications.

Research Method

For data gathering, website of Journal of Information Technology Management was used. Name of authors, their affiliations and quanitity of papers were extracted from this website. To calculating of Authors Collaboration Coefficient in Jouranl of Information Technology Management, this formula (Ajiferuke, Burell, and Jean Tague, 1998) was used.

$$cc = 1 - \left\{ \sum_{j=1}^{k} \left(\frac{1}{j} \right) * \frac{Fj}{N} \right\}$$

For extracting highly cited authors and the most active universities that have collaborated in publishing papers in Jouranl of Information Technology Management URL of ISC database (http://sci.isc.gov.ir/Search.aspx) was used.

Research Findings

The results revealed that the Collaboration Coefficient in Journal of Information Technology Management was 0.608 that describes a relatively suitable level. Findings indicated that Tehran University (with 71 articles), Allameh Tabatabee University (with 21 articles) and Tarbiat Modares University (with 14 articles) have had the most grouping published articles in Journal of Information Technology Management. In this journal, just 6 articles were individual and 152 articles were collaborative. Mohammad Musakhani with 7 articles was the most active author and Ali Asghar Anvari Rostami and Benam Shahaee with 8 citations were highly cited authors of current journal.

Table 1. Average of authors in Journal of Information Technology Management for each paper

Publication Year	Average of aurhors for each paper	Number of Authors	Number of papers
2009	2/3	19	8
2010	2/6	39	15
2011	4/65	93	19
2012	2/69	97	36
2013	2/91	105	36
2014	2/7	81	29
-	**3/01**	**476**	**158**

Table 1 shows that 158 papers were published in in Jouranl of Information Technology Management during 2009-2013 and 476 authors had publications in this journal. Most of papers were published in 2012 and 2013 years, 36 papers were published. In 2012, 105 authors published papers in this journal. Average of authors in 5 years was 3.01

Table 2. Collaboration of universities in studied papers

Name of universities	Number of collaborative papers
Tehran University	71
Allameh Tabatabee University	21
Tarbiat Modares University	14

Table 2 indicates scholars of Tehran University, Allameh Tabatabee University and Tarbiat Modares University had the most group papers in Jouranl of Information Technology Management. scholars of Tehran University with 71 group papers in Jouranl of Information Technology Management had the most collaboration in producing scientific papers.

Table 3. Frequency of studied papers on basis on number of each paper authors

Number of papers				Year
Four Authors	Three Authors	Two Authors	One author	
1	1	6	-	2009
2	5	8	-	2010
6	7	6	1	2011
6	15	14	1	2012
11	13	10	2	2013
14	12	15	2	2014
40	53	59	6	Total

Table 3 reveales that from 158 published papers in Information Technology Management, 6 papers were individual authors and 152 papers were group. This table shows that Iranian scholars trend to collaborative scientific productions and group papers had growing process.

Table 4. Authors Collaboration Coefficient in Jouranl of Information Technology Management during 2009-2014

Year	Authors Collaboration Coefficient
2009	0/56
2010	0/6
2011	0/62
2012	0/62
2013	0/63
2014	0/62
Average	**0/608**

Table 4 indicates the authors collaboration coefficient in Journal of Information Technology Management during 2009-2014. Author's collaboration coefficient is a number among 0 and 1. If this number is more than 0.5 means that collaboration between authors is in favorable level.

In table 4, it can be seen that authors collaboration coefficient in Jouranl of Information Technology Management is 0.608 that describe desireable level.

Table 5. Highly cited authors of Journal of Information Technology Management in Islamic Science Citation (ISC) database

Name of Authors	Number of citations
Ali Asghar Anvari Rostami, Behnam Shahaee	8
Farajollah Rahnavard, Asghar Mohammadi	6
Arian Gholopour, Behnam Amiri	5
Ahmad Roosta, Abalfaz Abalfazli, Hasan Ghorbani	4
Farajollah Rahnavard, Jalil Khavandkar	3
Maliheh Siavashi, Bahareh Abedin	3

Table 5 shows highly cited authors of Journal of Information Technology Management in Islamic Science Citation (ISC) database. It indicates Ali Asghar Anvari Rostami and Behnam Shahaee with 8 citations received the most citation.

Table 6. The most active authors in Journal of Information Technology Management during 2009-2014

Name of authors	Number of Papers
Mohammad Musakhani	7
Amir Manian	6
Shahriar Azizi	5
Hamid Reza Yazdani	4
Ali Mohammadi	4
Ali Mohaghar	4

Table 6 indicates that Mohammad Musakhani with 7 papers was the most active author in Journal of Information Technology Management. After he, Amir Manian and Shahriar Azizi with 6 and 5 papers in the second and third ranks.

Conclusion

Current research revealed that scholars had fewer trends to publish one author paper in Journal of Information Technology Management. 475 authors published 158 papers in journal of Information Technology Management during 2009-2014. Current research revealed that 6 papers were individual and 152 papers were group. In average for each paper, 3.01 authors had collaboration. Findings indicated that author's collaboration coefficient in Journal of Information Technology Management was 0.608 that means this is a desirable status. Current research revealed that papers of Journal of Information Technology Management that authors had trend to collaboration and group papers. Sharing in knowledge, resources and responsibilities are considered in most of scientific disciplines, so group works shape most of publication.

References

Ajiferuke, Isola, Q. Burell, and Jean Tague (1998). Collaborative coefficient: A single measure of the degree of collaboration in research. *Scientometrics* 14.5: 421-433.

Farajpahlou, A. Hossein (2004). Collaboration among Library and Information experts vs. scientist. *International Workshop on Webometrics, Infometrics and Scientometrics & 5th COLLNET Meeting, 2-5 May, Roorkee, India.*

Gupta, BM, Dhawan, SM (2007). Status of Physics Research in India: An Analysis of Research Output during 1993-01. *International Workshop on Webometrics, Infometrics and Scientometrics & 8 Th COLLNET Meeting, 10-12 May, 2007, India.*

Hart, Richard L. (2000). Collaborative publication by University librarians: An exploratory study. *Journal of Academic Librarianship*. 26(2): 94-99.

Zouhayr Hayati, Fereshteh Didegah, (2010) International scientific collaboration among Iranian researchers during 1998-2007, *Library Hi Tech*, Vol. 28 Iss: 3, pp.433 – 446.

Liang, LIMING, Kretschmer, Hildruan, Guo, Yongzhen, Beaver, Donald deb (2001). Age structures of scientific collaboration in Chinese computer science. *Scientometrics,* Vol. 52, No. 3: 471–486

Noruzi, Alireza, ALimohammadi Dariush (2006). Scientific Collaboration of the Iranian LIS Professionals across the World: With an Emphasis on Citation Indexes (1971-2006)**.** *Informology***.** Vol.3,No.3&4

Osareh,F;Marefat,R.(2005).collaboration of Iranian researchers in producing world science in Medline .*Rahyaft Quarterly*.35(spring):39-44.

Osareh, F.; Wilson, C. S. (2002). Collaboration in Iranian Scientific publication. *Libri*. 52(2): 88-98.

Sarrafzadeh, M. (2000). Reflection of Iranian articles in tow databases CAB and AGRIS from the first till 2000. *Rahyaft* Quarterly, 22 (Spring and Summer):88-97.

Osareh,F.(2005 a).Collaboration in Astronomy knowledge production: A case study in Science Direct from 2000-2004. Held in *10th International Conference on Scientometrics and Informetrics, 24-28 July in Stockholm-Sweden.*

Scientific Cooperation Engineering Making Interdisciplinary Knowledge Available within Research Facilities and to External Stakeholders

André Calero Valdez*, Anne Kathrin Schaar*, Tobias Vaegs**, Thomas Thiele**, Markus Kowalski**, Susanne Aghassi***, Ulrich Jansen****, Wolfgang Schulz****, Guenther Schuh***, Sabina Jeschke** and Martina Ziefle*

*Human-Computer Interaction Center, Campus Boulevard 57, RWTH Aachen University, Germany
calero-valdez@comm.rwth-aachen.de

**IMA/ZLW & IfU, Dennewartstr. 27, RWTH Aachen University, Germany

***Fraunhofer Institute for Production Technology, Aachen, Germany

****Department of Nonlinear Dynamics of Laser Processing, Steinbachstr. 15,
RWTH Aachen University, Germany

Abstract

In this paper we introduce the Scientific Cooperation Portal (SCP), a social enterprise software, and how it is integrated into our process of Scientific Cooperation Engineering. This process is applied in a large-scale interdisciplinary research cluster to ensure and manage the success of the interdisciplinary cooperation of over 180 researchers in different qualification levels. We investigate the influence of shared method competencies as an exemplary driver for collaboration. From the results we address both offline and online measures to improve interdisciplinary collaboration. We show how the knowledge generated from offline measures such as colloquia are transferred to the SCP and connected with other data available on the portal. This includes the handling of interdisciplinary terminologies, the disposability of publications and technology data sheets. The portal fosters knowledge exchange, and interdisciplinary awareness within the research cluster as well as technology dissemination both within the cluster, across the university, and into industry. The effectiveness of the approach is continuously assessed using a traditional balanced scorecard approach as well as additional qualitative measures such as interviews and focus groups.

Introduction

Dealing with complex global challenges often requires interdisciplinary research approaches to find suitable solutions (Repko 2012). Staying within disciplinary boundaries may prevent researchers to get a holistic overview of the topic at hand. Although the term interdisciplinarity lacks a unified definition (Jungert et al. 2010) it can be seen as the successful cooperation of researchers trained in the methods and conceptual approaches of different disciplines. Interdisciplinary research integrates these various methods to create new insights and methods for complex problems. Yet, actually making interdisciplinary research happen can be cumbersome because of lacking a common language, method competencies and understanding of scientific success. This problem intensifies under conditions of high staff turnover, research group size (Repko 2012), performance pressure, and increasing complexity of the research problem. How to measure interdisciplinary collaboration and finding reasons for this collaboration, and the deliberate steering of interdisciplinary groups are still largely unsolved questions. Thus active support for such collaboration requires various measures and a constant evaluation of these measures. We apply findings from bibliometrics and cybernetics to management principles of a research cluster in order support interdisciplinary collaboration and scientific success of the cluster.

Related work

Collaboration trumps solo-efforts in generating knowledge (Wuchty et al. 2007). Finding evidence of (interdisciplinary) collaboration can traditionally be done by analyzing co-authorship networks (Glänzel & Schubert 2005), although one must be careful not to mistake co-authorship for collaboration and vice versa (Melin & Persson 1996). Investigating who publishes with whom can reveal collaboration patterns and thus be used to understand interdisciplinary cooperation. Glänzel & Schubert found that geopolitical location and language are determining factors for collaboration. Collaboration decreases exponentially with physical distances (Katz 1994, Hoekman et al. 2010). Kretschmer (1999) found that similarity as well complementarity can be used to explain researchers' collaboration by analyzing co-authorship relationships. By applying this approach Kretschmer & Kretschmer (2012) could explain up to 99% of the variance for 77% of the co-authorship relationships. De Solla Price & Gürsey (1975) identified different types of authors according to their publishing behavior (i.e. continuants, transients, recruits, terminators) for which Braun et al. (2001) identified differing author productivity and collaboration patterns. Newman (2001) found patterns of small world phenomena (i.e. short paths between any two random authors). Co-author networks showed various levels of clustering and a fractal nature (e.g. self-similarity). Van Raan (2000) developed a model to determine growth of scientific literature based on the fractal nature of science. Sub-systems grow individually and can be seen as self-organizing units. This reflects in the cybernetic nature of how universities are managed (see Birnbaum & Edelson 1989). Cybernetics in this regard means that no centralized "premeditated" plan (for publications) is conceived by the management but, in the manner of a thermostat, a target output is defined and measures are taken to reach the target.

Using interviews Hara et al (2004) created a model for determining factors of collaboration in in a research center. From the interviews they found two different types of collaboration, "complementary" and "integrative" collaboration. Determining factors were compatibility (i.e. work style, priority, management style, approach to science, personality), work connections (i.e. work interests, expertise), incentives (i.e. external funding, publication, internal) and socio-technical infrastructure (i.e. awareness, communication mechanism, organization culture and structure, access to collaborators). Overall they assume personal relationships beget professional relationships and thus collaboration. They suggest that technological support could enhance the process of collaboration and that it needs further investigation.

Various forms of these collaboration support systems exist. This new emerging field of E-Science and E-Infrastructure draws on the tools and methods developed from Computer-Supported Cooperative Work (Jirotka 2012). Zheng et al. (2011) present TSEP a social platform to assist collaboration between scientists. Li et al. (2012) and Müller-Tomfelde et al. (2011) strengthen the need for shared workspaces and audio-visual support of workgroups in a health laboratory, but also tailoring to the needs of the workgroup. Alves et al. (2013) have suggested a system for finding possible collaborators in a scientific setting. Romano et al. (2011) suggest the use of wikis and ontologies along with learning environments to support researchers in the field of bioinformatics. Above all tailoring a Social-Network-Solution (SNS) to the users needs is critical, as communicative preferences may depend on user characteristics (Calero Valdez et al. 2012a).

Research Questions

In this paper we demonstrate the efforts undertaken in a research cluster to support interdisciplinary collaboration. For this purpose we look into both online and offline measures that support collaboration. We assume that shared method competencies may also be a driver of collaboration. Here we compare the shared method competencies of workgroups generated from both publication data and qualitative data collected at a member colloquium. Furthermore we show how the insights from the study are used as feedback to the researchers in the cluster. In the following sections we first describe the research cluster, the Scientific Cooperation Portal and then the analysis of methods used in the cluster.

The Scenario - The Aachen Cluster of Excellence

The challenge of keeping production industry sustainable in countries with high wages is also in interdisciplinary one. In the research cluster of excellence (CoE) "Aachen House of Integrated Production" researchers from various subfields of physics, material sciences, engineering, computer science, up to economics and social sciences are faced with the challenges of production on various levels of scale and their interfaces (i.e. from raw material properties to production processes to factory and logistics planning, with respect to human needs on all of these levels). Overcoming the stereotypic scale-scope dilemma (individualized products vs. mass production) of production (Brecher 2012) is one key goal of this research cluster. Additionally it faces the unification of the dilemma of plan- vs. value-oriented production, in conjunction called the polylemma of production. In total about 180 researchers work on this holistic view on production technology, grouped in different working areas. These researchers work in four integrated cluster domains (ICDs), which are interconnected by so called cross-sectional processes (CSPs, see Figure 1). These CSPs ensure sustainability of the research cluster in regard to human resources, advancement of scientific theory and development of technology platforms (Jooß 2012). Their research goal is to investigate, what methods work effectively to achieve said sustainability. Additionally they assist the steering committee of the cluster by providing insights on performance and recommending a course of action.

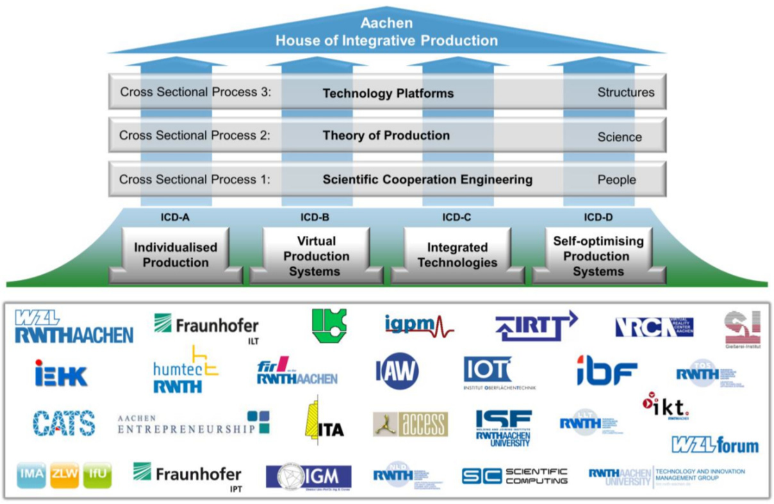


Figure 1. Research structure of the CoE, integrating institutes from five faculties of RWTH Aachen University and focusing on sustainability within the dimensions people, science and structure, incorporated within the Aachen House of Integrative Production (Brecher 2012).

Managing Collaboration

In order to ensure that the cluster works effectively key performance indicators (KPI) are established to measure performance for both internal (management) and external use (funding agency evaluation). This is done using a balanced-score-card approach (Welter 2011) with typical performance measures as (peer-reviewed) publications, patents and third-party funding, but are also contrasted by criteria like knowledge dissemination, interdisciplinarity, quality of supervision, and many more. These are used to determine how well the cluster works and where it needs improvement.

Bringing researchers from so many scientific fields together requires management of many of these success criteria in an individualized fashion. Disciplines differ in regard to what is considered successful as a publication or as advancement in theory. In order to unify the dilemma of required disciplinary diversity and the need for a unified measure of success a cybernetic management approach is applied. For example, indicators are developed that measure the transfer of knowledge within the cluster, the development of interdisciplinary methods, the coherence of the research road map, or the transfer of technology within the cluster and into industry.

Measuring performance in an interdisciplinary context is not a trivial task, but beyond that, steering performance is even harder. The cybernetic management approach incorporates various measures to both measure and steer performance.

A mix of offline and online measures is used to reach a maximum of potential cluster members. As offline steering measures the CSPs conduct member colloquia, cluster conferences, general assemblies, seminars, and workshops. In the member colloquia all partaking researchers spend a whole day dealing with topics that overarch the ICD-structure of the cluster, such as interdisciplinary communication skills (e.g. presenting research to non-experts), finding research partners (e.g. scientific speed dating) and developing a common research road map. On dedicated cluster conferences researchers present the results of their individual scientific research to the other members. In general assemblies principle investigators (PI) present the meta-level of research from their institutional point of view connecting the theory behind partaking institutes. These measures foster the interdisciplinary awareness, cooperation, communication and method skills. Some topics are addressed in seminars or workshop to address individual and sub-project based needs. For example a seminar on interdisciplinary publishing addresses the participants perception of the publishing process form their disciplinary perspective. Best-practices in cluster-typical cooperation are discussed and shared with the participants. An online method to enrich these offline approaches is the Scientific Cooperation Portal presented in this paper.

All measures are all evaluated in regard to the KPIs quantitatively (using a questionnaire method) but they are also addressed in interviews and focus groups with the researchers to ensure validity of the measurements.

The Scientific Cooperation Portal

As an online measure the CSPs introduced the Scientific Cooperation Portal (SCP) in 2013 (Vaegs 2014). The SCP is a social portal system used as a centralized knowledge storage system and was introduced to face the aspect of transparency of communication, which appeared in several evaluations. Voluntary access to the SCP is limited to cluster members and PIs exclusively (yet).

The SCP provides user profiles, yellow pages, a cluster based news feed, calendar and event system, and a centralized file storage system. Required forms for typical needs (e.g. travel expense forms) are available from this centralized storage system. All data on the SCP can be tagged and thus interconnected with each other. As specific features designed to match the cluster specific needs measured by the BSC, interviews, and focus groups, applications are built to address the challenges of interdisciplinary use of terminology, interdisciplinary publications, and technology transfer.

User Profiles

Members profiles can be found through the yellow page system and contain information about disciplinary background, method competencies, expertise in technology, publications, and participation on terminology definitions. Furthermore typical contact information is available.

Terminologies

One critical aspect mention in many evaluations is the lack of a unified language/terminology. Since different disciplines use terminology differently the approach of the CSPs is not to unify terminology, but to enhance awareness of disciplinary differences. For this purpose an application is developed that portrays the differing definitions of frequently used terms from the various perspectives, highlighting differences in understanding. Definitions are connected to their authors, publications in which they are used, and their technology data.

Publication Relationship Analysis

Publications are a peculiar aspect of scientific work, as they disseminate knowledge gain to the scientific community. They are often (wrongly) used as sole performance indicators overvaluing quantity above quality. The SCP uses publications to establish researcher profiles. This allows the CSPs to understand (and measure by proxy) the collaboration in the CoE. Furthermore we will use visualization and graph based approaches to understand and communicate publishing efforts of the CoE to its members (Calero Valdez 2012b). User profile pages will be connected with their co-authors, but also with topics stemming for publications keywords. Furthermore used technology and terminology from publications are connected with their respective technology data sheets and terminology pages.

Technology Transfer

Technology developed in the CoE should be disseminated both within and to industry partners to be useful to a possible consumer of the technology. In order to simplify communication of advances, a technology transfer portal is integrated into the SCP (Schuh 2013). Here technology data sheets present key advantages of developed technology and contact information of the provider of the technology (see Figure 2). They are also connected to their provider users as well as publications that relate to the technology. Technology data sheets can be customized to be viewable by external partners (e.g. industry) once they have achieved a sufficient level of stability.

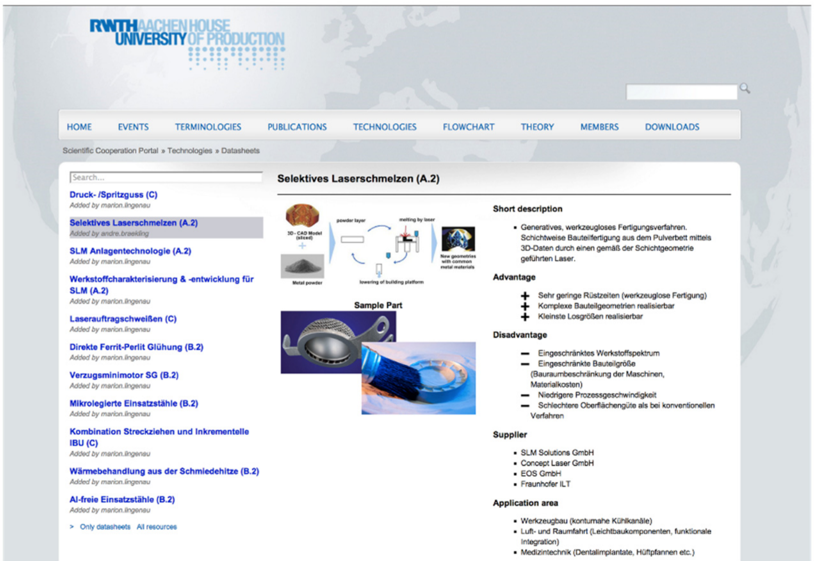


Figure 2. Example technology data sheet on the SCP.

Methodology – Assessing Method Competencies

In order to find out what methods are used in the cluster we approach that topic from two directions. First we pick full-text data from the cluster and manually scan the methodology sections of these papers for named-entities that refer to method-names. We then perform manual deletion of duplicates on synonyms on the data. We create a method graph connecting each workgroup with its methods. Since classical database coverage of engineering sciences is subpar (Harzing & Van der Wal 2007), we collect publication data manually by requiring researchers to submit their work in order collect funding for travel expenses for instance.

In a second step, conducted during a member colloquium, we asked all workgroups to brainstorm on the methods that they used on a daily basis (see Figure 3). The time frame for this task was about 90 minutes, and instructions were given to collect methods that are both used in publications and methods that are available but have not been used yet. As a working definition what constitutes a method several definitions were given (US patent definition, a definition derived from philosophy of science, a definition from Computer Science) to heighten awareness of disciplinary differences in the meaning of the term "method". Methods are then again cleared for duplicates and synonyms. Another method graph is constructed. Both method graphs are then compared an evaluated in regard to graph statistics.

Not addressed in this paper are the workshops that address in a similar fashion the topics of interdisciplinary terminologies and technology data sheets.

Figure 3. Exemplary results of a method workshop in a subproject.

Results and Interpretation

At this current timeframe full-text publications were available for 7 of 12 sub-projects. From over 500 publications 76 were selected (availability and containing a clear method section) and manually scanned for methods. Form these, 222 named-entities were recognized and reduced to 195 unique methods. The constructed method graph (see Figure 4) showed a graph density of *.006*. Community detection (Blondel et al. 2008) revealed 7 communities and a modularity of *.773*.

The method collection from the member colloquia surprisingly also resulted in a sum of 195 methods (after deletion of duplicates and synonyms). The graph (see Figure 4) showed a graph density of *.005* and also revealed 7 communities. Modularity of the graph was determined at *.766*.

Interestingly the nodes connecting most sub-projects in both graphs are nodes that relate to "modelling", "FEM" and "Software Development". Method overlap in both cases is sparse, meaning that either shared methods are sparse, remain unmentioned (in both verbal an written communication) or that no unified terminology exists regarding applied methods. Both graphs show a structural symmetry between each other.

As a side note is worth mentioning that even the term "method" is far from having a shared understanding. During the member colloquium the need for clarification arose, in particular in regard to discerning it from the term "technology". In the various fields of engineering, clear differentiation is not always possible. One develops a technology that is used by others as a method. Discussions regarding this took substantial time off of brainstorming times.

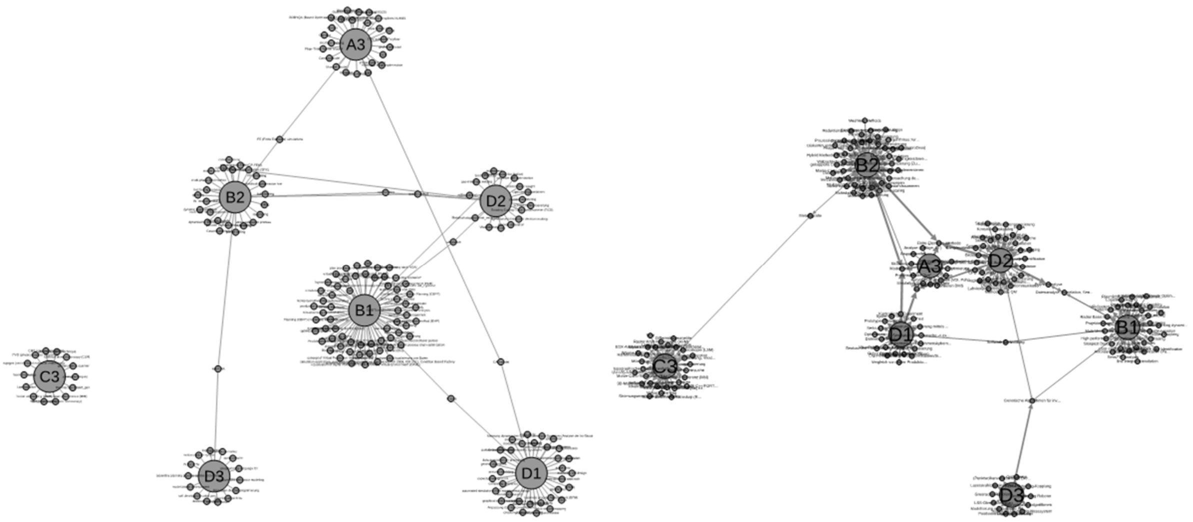


Figure 4. Method graphs constructed from member colloquium data (left) and from method sections of publications (right)

Conclusion

The differences in terminology, in particular in regard to the term "method" itself, further underline the need for support in an interdisciplinary setting. As mentioned by Hara et al. (2003) compatibility is essential for scientific collaboration.

Applying the approach from Alves et al. (2013), we enrich researcher's profiles with method competencies to enable finding researchers within the cluster that share research interests. The terminology application must respect disciplinary differences in understanding of methods (that can also be technologies) and can be seen as a measure to broaden understanding of method competencies across disciplinary boarders. Furthermore technology transfer must be performed not only to external stakeholders but also within a research cluster. The findings from the member colloquium confirm the need for social software that integrates terminology, methodology, technology, and publications as an online support measure to our research cluster. This means when a user opens another user's profile, he will see a list of methods used by this researcher, which hyperlinks to an ontology-based wiki and also full-text publications (when available) that contain these methods. Furthermore technology used by a researcher is hyperlinked to technology data sheets, which in turn are linked to publications and terminology.

In the future, we are able to better understand interdisciplinary cooperation by following the individual as well as the work groups' usage behavior of information of the Portal. Both, the genesis of a novel cooperation can be retraced and related to the respective genesis conditions as well as the growing density of the collaboration's network in order to see growing novel topics or methodologies within and across work groups. Also, looking from the industry side and the analysis of industry's interest and search for information behavior can be also a promising approach for emerging topics and research fields.

Limitations

The procedures to generate graphs rely heavily on manual correction and synonym detection. We must assume that further unnoticed synonyms exist in the data as the author is no expert in all of the found methods. This limitation also applies to the manual named-entity search in the papers. Furthermore only a fraction of the actual publication output was used, due to availability of full texts.

The similarity of the graph could to a large extend be caused by the method of construction. For both graphs first workgroup nodes are created and then connected to their method nodes. This would in many cases lead to similar graphs, if methods were unrelated.

The presented approach was used as a starting point into the data. In the future users of the portal may choose to add their own synonyms to method definitions to enhance the analysis process in future iterations. The approach also only reflects collaboration of the similarity type. Complementary or integrative collaboration should in essence not contain the same set of methods. Nonetheless an overlap that enables communication should be found.

Furthermore we have not looked into interrelations between both graphs yet, as the methods are not in a single language. Finding adequate translations should also be a user driven task as well.

Summary and Outlook

In this paper we presented the scientific cooperation portal a social portal to support interdisciplinary collaboration in research clusters. The features of the portal were developed from systematic evaluation of researchers needs using both qualitative and quantitative methods (Schaar 2013). Content for the portal is generated by both the users and the CSPS from at various events. Furthermore we looked into shared method competencies as a driver for collaboration by investigating the methods used in the sub-projects both from verbal and written evidence. We found low overlap between sub-projects in methods, but high similarity for both approaches. Interestingly when comparing the method overlap with actual collaboration from publication data (Calero Valdez et al. 2012b), we find a similar graph density (.005) but a higher level of clustering (27 communities, modularity .844). Further evaluation (e.g. graph isomorphism) will reveal whether this accurately reflects similarity between the different graphs. Furthermore looking into references and citation data could prove useful. Researchers sharing the same methodology should cite similar work. The hypothesis that ones technology is another's method could also be verified by looking into citations in method sections. From these findings we derive the need for collaboration support and underline the selection of features of the Scientific Cooperation Portal as well as conducting member colloquia which bring researchers together on a personal level and foster communication between sub-projects and across disciplinary boarders.

Connecting both offline with online measures has improved KPIs for scientific collaboration, which was established by a BSC-approach.

Acknowledgements
We would like to thank the participants in the colloquium for the efforts. We would also like to thank Tatjana Hamann and Juliana Brell for their support. This research was funded by the Excellence Initiative of the German federal and state governments.

References

Alves, T. P., Borges, M. R., & Vivacqua, A. S. (2013, June). An environment to support the discovery of potential partners in a research group. In *Computer Supported Cooperative Work in Design (CSCWD), 2013 IEEE 17th International Conference on* (pp. 344-349). IEEE.

Blondel, V. D., Guillaume, J. L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, *2008*(10), P10008.

Birnbaum, R., & Edelson, P. J. (1989). How colleges work: The cybernetics of academic organization and leadership. *The Journal of Continuing Higher Education*, *37*(3), 27-29.

Braun, T., Glänzel, W., & Schubert, A. (2001). Publication and cooperation patterns of the authors of neuroscience journals. *Scientometrics*, *50*(3), 499-510.

Brecher, C., Jeschke, S., Schuh, G., Aghassi, S., Arnoscht, J., Bauhoff, F., & Welter, F. (2012). *Integrative production technology for high-wage countries* (pp. 17-76). Springer Berlin Heidelberg.

Calero Valdez, A., Schaar, A. K., Ziefle, M., Holzinger, A., Jeschke, S., & Brecher, C. (2012b). Using mixed node publication network graphs for analyzing success in interdisciplinary teams. In *Active Media Technology* (pp. 606-617). Springer Berlin Heidelberg.

Calero Valdez, A. , Schaar, A. K., & Ziefle, M. (2012a). State of the (net) work address Developing criteria for applying social networking to the work environment. *Work: A Journal of Prevention, Assessment and Rehabilitation*, *41*, 3459-3467.

Cummings, J. N., & Kiesler, S. (2005). Collaborative research across disciplinary and organizational boundaries. Social Studies of Science, 35(5), 703-722.

Glänzel, W., & Schubert, A. (2005). Analysing scientific networks through co-authorship. In *Handbook of quantitative science and technology research* (pp. 257-276). Springer Netherlands.

Hara, N., Solomon, P., Kim, S. L., & Sonnenwald, D. H. (2003). An emerging view of scientific collaboration: scientists' perspectives on collaboration and factors that impact collaboration. *Journal of the American Society for Information Science and Technology*, *54*(10), 952-965.

Harzing, A. W., & Van der Wal, R. (2007). Google Scholar: the democratization of citation analysis. *Ethics in science and environmental politics*, *8*(1), 61-73.

Hoekman, J., Frenken, K., & Tijssen, R. J. (2010). Research collaboration at a distance: Changing spatial patterns of scientific collaboration within Europe.*Research Policy*, *39*(5), 662-673.

Holzinger, A., Ofner, B., Stocker, C., Valdez, A. C., Schaar, A. K., Ziefle, M., & Dehmer, M. (2013). On Graph Entropy Measures for Knowledge Discovery from Publication Network Data. In *Availability, Reliability, and Security in Information Systems and HCI* (pp. 354-362). Springer Berlin Heidelberg.

Jirotka, M., Lee, C. P., & Olson, G. M. (2013). Supporting Scientific Collaboration: Methods, Tools and Concepts. *Computer Supported Cooperative Work (CSCW)*, *22*(4-6), 667-715.

Jooß, C., Welter, F., Leisten, I., Richert, A., Schaar, A., Valdez, A. C., & Jeschke, S. (2012). Scientific cooperation engineering in the cluster of excellence integrative production technology for high-wage countries at RWTH Aachen University. *ICERI2012 Proceedings*, 3842-3846.

Jungert, M., Romfeld, E., Sukopp, T., and Voigt, U. (2010). *Interdisziplinarität: Theorie, Praxis, Probleme*, 1st ed. Darmstadt, Hesse: Wissenschaftliche Buchgesellschaft, 2010.

Katz, J. S. (1994). Geographical proximity and scientific collaboration. *Scientometrics*, *31*(1), 31-43.

Kretschmer, H. (1999). A new model of scientific colloboration part 1. Theoretical approach. Scientometrics, 46(3), 501-518.

Kretschmer, H., & Kretschmer, T. (2012). Who is collaborating with whom in science? Explanation of a fundamental principle. *Social Networking*, Scientific Research Publishing.

Li, J., Muller-Tomfelde, C., & Robertson, T. (2012, January). Designing for distributed scientific collaboration: a case study in an animal health laboratory. In *System Science (HICSS), 2012 45th Hawaii International Conference on* (pp. 373-381). IEEE.

Melin, G., & Persson, O. (1996). Studying research collaboration using co-authorships. *Scientometrics*, *36*(3), 363-377.

Müller-Tomfelde, C., Li, J., & Hyatt, A. (2011). An integrated communication and collaboration platform for distributed scientific workgroups. In *Human-Computer Interaction–INTERACT 2011* (pp. 248-258). Springer Berlin Heidelberg.

Newman, M. E. (2001). The structure of scientific collaboration networks.Proceedings of the National Academy of Sciences, 98(2), 404-409.

van Raan, A. F. (2000). On growth, ageing, and fractal differentiation of science. *Scientometrics*, *47*(2), 347-362.

Repko, F. (2012). *Interdisciplinary Research: Process and Theory*. Los Angeles, CA: Sage.

Romano, P., Giugno, R., & Pulvirenti, A. (2011). Tools and collaborative environments for bioinformatics research. *Briefings in bioinformatics*, *12*(6), 549-561.

Schaar, A. K., Valdez, A. C., & Ziefle, M. (2013). Publication network visualization as an approach for interdisciplinary innovation management. In *Professional Communication Conference (IPCC)*, 2013 IEEE International (pp. 1-8). IEEE.

Schuh, G., Aghassi, S., & Valdez, A. C. (2013). Supporting technology transfer via web-based platforms. In *Technology Management in the IT-Driven Services (PICMET)*, 2013 Proceedings of PICMET'13: (pp. 858-866). IEEE.

de Solla Price, D., & Gürsey, S. (1975). Studies in Scientometrics I Transience and Continuance in Scientific Authorship. *Ci. Inf*, *4*(1), 27-40.

Vaegs, T. et al. (2014 in press). Enhancing Scientific Cooperation of an Interdisciplinary Cluster of Excellence via a Scientific Cooperation Portal. Full paper accepted at the International Conference on E-Learning at the workplace, June 11-13, Columbia University, New York

Welter, F., Vossen, R., Richert, A., & Isenhardt, I. (2011). Network Management for Clusters of Excellence-A Balanced-Scorecard Approach as a Performance Measurement Tool. In *Automation, Communication and Cybernetics in Science and Engineering 2009/2010* (pp. 195-207). Springer Berlin Heidelberg.

Wuchty, S., Jones, B. F., & Uzzi, B. (2007). The increasing dominance of teams in production of knowledge. *Science*, *316*(5827), 1036-1039.

Zheng, X., Ke, G., Zeng, D. D., Ram, S., & Lu, H. (2011). Next-generation team-science platform for scientific collaboration. *Intelligent Systems, IEEE*, *26*(6), 72-76.

Measuring Efficiency of Scientific Research

Arshia Kaul, Sujit Bhattacharya*, Shilpa and Praveen Sharma

CSIR-National Institute of Science Technology and Development Studies
Dr. K.S. Krishnan Marg, New Delhi-110012, India
**sujit.personal@gmail.com*

Introduction

Measuring science and technology outputs in relation to the inputs helps to measure the efficiency of the scientific system. Outputs in scientific research are both tangible and intangible. Indicators can be constructed from tangible outputs to provide a 'proxy' measure of the various components of the system. In case of S&T enterprise, indicators are constructed from research publications and various types of translational research to assess the performance of scientific system. The translational research outcomes can comprise varied types of technological outputs such patents, products, processes, instruments and designs. The intangible outputs are tacit in nature and difficult to codify and thus quantifiable measurement based indicators is difficult to construct from them.

A large volume of research is available that discusses the various indicators that measure science and technology (see for example Godin 2004). Increasingly it is becoming essential to construct indicators that capture the input-output together to have a better representation of the overall system. The knowledge gained from analyzing critically the amount of output as a result of the input received for research (i.e. return to investment) helps the decision maker to make policy decision.

In this paper we have constructed two indicators to capture the efficiency of S&T enterprise. The S&T enterprise chosen for investigation was CEFIPRA (Indo-French Centre for Promotion of Advanced Research), a bilateral centre that was set-up between India and France in 1987 with the objective to promote STI (science technology and innovation) cooperation between the two countries. The centre helps to promote collaboration in fundamental and applied scientific research, identify scientists and scientific institutions, assistance in terms of grants and equipment as well as other appropriate means for pursuing advanced research, organisation of workshops or seminars and other types of activities in areas of mutual interest.

Scientific projects funded by this institute for the last 25 years from 1987 to 2012 was object of this investigation. The two indicators were constructed to measure the efficiency of the projects funded by this organisation.

Construction of Indicators

Two indicators that were constructed are: (a) Technology Output Index, and (b) Volume-Value Index. For analysis and construction of two indicators the projects have been allocated to three quartiles as High, Medium and Low based on the number of projects in a Thrust area as given in Table 1 (CEFIPRA has identified 12 areas in which it gives funding as Thrust areas. Please note that High, Medium, and Low that the study undertook to classify the thrust areas are not in terms of significance of the project or Thrust area).

Technology Output Index

Technology Output Index was constructed to properly weight technological output produced in each thrust area from the projects. Technology outputs covered are processes created, products developed, patents filed, design created, and instrument developed. Equal weight of 1/5 is applied to the frequency of occurrence of each technology output in a Thrust area to calculate the overall composite index. If a Thrust area contributed to the development of 37 processes, 2 products, 2 patents and there are no designs or instruments developed from any of the projects in that Thrust Area, then the index value for that Thrust area was calculated as (0.2*37+0.2*2+0.2*2=8.2). The value for each thrust area was calculated similarly.

Table 1 Quartile-wise Distribution of Thrust Areas

High Quartile	**No. of Projects**	**Patents Filed**	**Product**	**Process**	**Design**	**Instrument**	**Index Value**
Material Sciences	69	11	7	24	8	1	10.2
Life & health Sciences	75	4	6	37	1	1	9.8
Pure & Applied Chemistry	52	4	4	23	6	0	7.4
Pure & Applied Physics	51	3	2	22	3	0	6
Medium Quartile		**Patents filed**	**Product**	**Process**	**Design**	**Instrument**	**Index Value**
Earth & Planetary Sciences	15	0	0	11	3	0	2.8
Water	8	2	1	6	0	0	1.8
Biotechnology	8	0	0	6	2	0	1.6
Pure & Applied Mathematics	22	0	2	6	0	0	1.6
Computers & Information Sciences	14	0	0	4	2	1	1.4
Environmental Sciences	12	0	0	3	1	0	0.8
Low Quartile		**Patents filed**	**Product**	**Process**	**Design**	**Instrument**	**Index Value**
Instrumentation	3	0	0	0	1	0	0.2
TIC	4	0	0	1	0	0	0.2

As evident from Table 1 there has been maximum innovative output in Material Sciences (index value =10.2), followed by 'Life and Health Science' (index value =9.8) indicating the maximum contribution to innovation from high quartile. The trend roughly shows that the volume of the output is in relation with the number of projects. It is notable to observe that the thrust area 'Life and Health Sciences' has lower value of index as compared to index value for 'Material Sciences' although 'Life and Health Sciences' has higher number of projects. When an observation is made between the quartiles it shows that the high quartile as a group has a higher volume of output followed by the medium and low quartiles.

Volume-Value Index

The index constructed is named the Volume-Value Index. The index gives a relationship between the input variable (defined as Volume) and the output variable (defined as Value) for each thrust area. This relationship helps us to understand the efficiency of the scientific research system.

Volume and Value were defined as follows: Volume (in terms of input) was captured by three parameters: Number of projects funded by CEFIPRA in each Thrust area (P); Human Resource (HR) involved in each Thrust area, and Expenditure in a Thrust area (E). Each of the three parameters were given weights. Value was captured by two parameters Value I and Value II. Value I indicates research papers published from CEFIPRA funded projects in each Thrust area. Value II indicates technological output (processes created, products developed, patents filed, design created, and instrument developed) coming from each Thrust area. We examined the combined impact of Volume (P), Volume (HR) and Volume (E) on Value I and Value II.

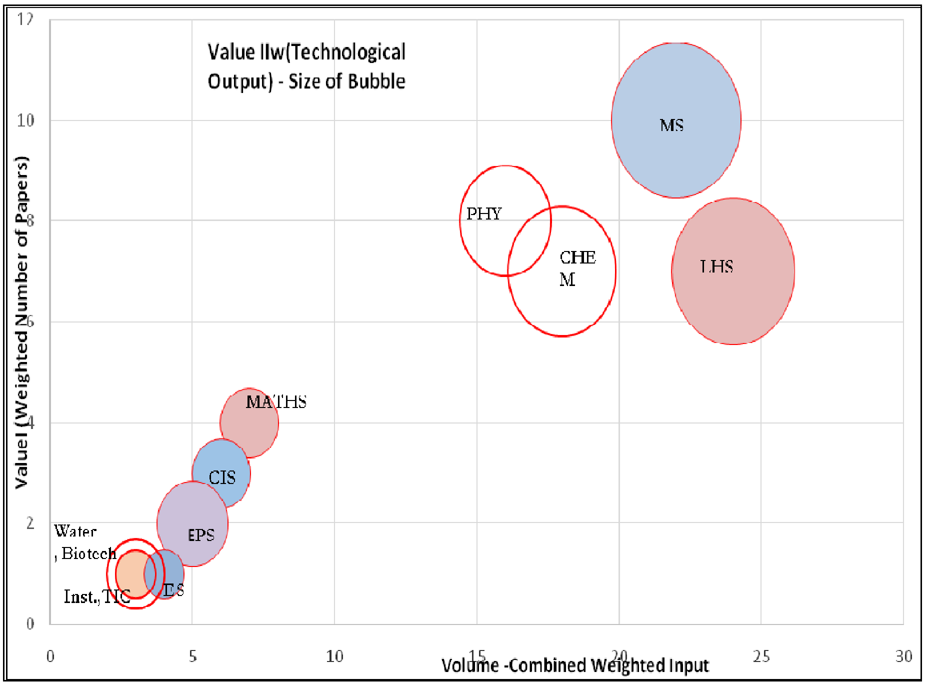


Figure 1. Volume vs. Value Analysis for Thrust Areas

Fig. 1 plots the weighted input Volume (w) [consisting of combined Volume (P), Volume (HR) and Volume (E)] with respect to Value I (weighted number of research publications) and Value II(weighted Technological Output). Fig.1 shows number of research papers increasing as Volume increases. A high correlation of 0.94 between Volume and Research papers (Value I) testifies to this fact. However, 'Material science' (MS) is not following the trend, scoring higher in terms of research papers than 'Life and health sciences' (LHS). Physics is also scoring just above Chemistry, slight deviation from the trend. *But the overall trend and high correlation implies higher the combined input leads to higher number of research paper productivity.* Thus, Thrust areas which have large number of projects, high expenditure, and human resources involved leads to high research paper output.

The correlation between Volume and Value II (weighted Technological Output) is much higher of 0.96. From the figure one can observe deviation of Water and Biotech from the trend i.e. higher the input, higher the Technological output. Earth & Planetary Sciences (EPS) also shows slight deviation. Thus from the Figure and statistical analysis *one can say that higher input leads to higher value. Also one can observe 'Material science' (MS) is the best performing Thrust area.*

Discussion and Conclusion

A scientific research can give rise to many types of outputs. The system needs to be analysed such that the efficiency of the scientific system can be adjudged.

A system can be considered as an efficient system if the output is high with less input. In this paper the focus has been to capture the efficiency of scientific research projects by constructing indicator which are a combination of the outputs and the inputs. In this context we understand the efficiency of various thrust areas of the CEFIPRA supported projects. The future scope of the research is to understand how to improve the efficiency different thrust areas. This can be done using the data envelopment analysis in which one is able to maximise the efficiency of one system with respect to other system given a certain inputs.

The indicators as discussed in this paper can be used to analyse the efficiency of the scientific research system of a country. Thus based on indications from such indicators policy makers will be able to improve efficiency in the system.

Acknowledgements

The authors thank the IFCPAR/CEFIPRA (Indo-French Centre for Promotion of Advanced Research/Centre Franco-Indien pour la Promotion de la Recherche Avancée) for supporting the present study.

Key Reference

Godin, B. (2004). Measurement and Statistics on Science and Technology: 1920 to the Present. Routledge.

How do scientists collaborate? Assessing the impact of influencing factors

Ashkan Ebadi and Andrea Schiffauerova

Concordia Institute for Information Systems Engineering (CIISE), Concordia University,
1515 Ste-Catherine Street West, Montreal, Quebec, H3G 2W1, Canada
a_ebad@encs.concordia.ca, andrea@ciise.concordia.ca

Abstract

Scientific collaboration is one of the substantial drivers of research progress that may lead researchers to generate novel ideas. Scientists may present such new thoughts in high quality journal publications or in the form of technology advances. There are several studies that examined collaboration networks or impact of network variables on scientific activities. However, to our knowledge this paper is the first that analyzes the impact of other influencing factors on network structure variables at the individual level. For this purpose, we focus on the collaboration network among funded researchers during the period of 1996 to 2010 and employ time related statistical models to estimate the impact on network structure variables. Results highlight the crucial role of past productivity of the researchers along with their available funding in determining and improving their position in the co-authorship network. It is shown that local influencers who possess high closeness centrality are not necessarily prolific researchers in terms of the quality of their publications. However, high quality productive researchers have higher betweenness centrality. Moreover, although mid-career scientists have higher closeness centrality, the role of young gatekeepers is confirmed in connecting different communities and information spread.

Keywords: network structure, collaboration, statistical analysis, Canada

Introduction

Thanks to the recent progress in information technologies nowadays no specific border can be defined for scientific activities in a way that researchers have formed a global community aiming to advance the level of knowledge. Concurrently, the nature of the science has become more complex and inter-disciplinary that encourages scientists to be more collaborative in an aim to increase their scientific productivity. However, collaboration may not necessarily augment the scientific performance and several issues need to be considered, e.g. selecting the right partner, coordination costs. Katz and Martin (1997) define scientific collaboration as the process through which the researchers with a common goal work together to produce new scientific knowledge. Scientific collaboration has been studied in a vast number of different disciplines such as computer science, sociology, research policy, and philosophy (Sonnenwald, 2007). Thourgh collaboration researchers get access to an often informal network of scientists that may facilitate knowledge and skill diffusion (Tijssen, van Leeuwen, & Korevaar, 1996; Tijssen, 2004). Although it is not easy to quantify scientific collaboration, co-authorship has become the standard way of measuring collaboration since it is considered as a better sign of mutual scientific activity (De Solla Price, 1963; Ubfal & Maffioli, 2011).

The importance of collaborative research is now acknowledged in scientific communities (Brad Wray, 2006), where financial investment can change the structure of research groups and affect the collaboration among the scientists. However, there might be some conflicts between individual preferences and the society level goals. These conflicts may cause different optimal individual collaboration level from the optimal social one (Ubfal & Maffioli, 2011). Although governmental funding for knowledge creation and diffusion has a long history, its effects on scientific collaboration and formation of scientific networks is relatively new (Katz & Martin, 1997; Lee & Bozeman, 2005). Researchers have started evaluating the impact of funding on the collaboration using simple indicators in the early 80s (e.g. Beaver & Rosen, 1979; Heffner, 1981).

Using econometric techniques and statistical analyses in some cases, a few studies recently assessed the impact of funding and other influencing factors like gender, past productivity, *etc.* on collaboration. Although some studies found a positive relation between funding and the scientific collaboration (e.g. Adams et al., 2005; Bozeman & Corley, 2004; Defazio, Lockett, & Wright, 2009; Gulbrandsen & Smeby, 2005), there also exits few studies that could not find any significant relation between funding and collaboration (e.g. Rosenweig et al., 2008). This study extends the literature in two ways. To our knowledge, no study has examined the impact of a group of influencing factors on the individual indicators of the position of researchers within their scientific collaboration network. In addition, most of the studies used a limited dataset and/or focused on a limited scope while this study uses a large dataset of NSERC[1] funded researchers. Our basic motivating questions are: How the influencing factors including funding affect the position of the scientists among their collaboration network? And, what are the most determinant factors in stimulating scientific collaboration? The remainder of the paper proceeds as follows: Section "*Data and Methodology*" presents the data, methodology and the model; Section "*Results*" presents the empirical results and interpretations; Section "*Conclusion*" concludes and suggests some directions for the future work.

Data and Methodology

Data

The data of this research was gathered in three phases. In the first phase, the funded researchers' data was collected from NSERC and then using Elsevier's Scopus[2] we gathered all the information (e.g. co-authors, their affiliations, year of publication) about the articles that were published by the funded researchers within the period of 1996 to 2010. We focused on NSERC since it is the main federal funding organization in Canada, and almost all the Canadian researchers in natural sciences and engineering receive a research grant from NSERC (Godin, 2003). We selected the period of 1996 to 2010 since the data quality of Scopus was low before 1996. To have a proxy for the quality of the papers we used SCImago[3] to collect the impact factor information of the journals in which the articles were published in. We chose SCImago since it provides yearly data of the journal impact factors that enables us to perform a more accurate analysis. In addition, SCImago is powered by Scopus that makes it more compatible with our articles database. In the second phase, we did a full text search over the articles and fetch the ones that acknowledged NSERC support in the body of the articles. This was a crucial step in gathering more accurate data since the common procedure in the similar studies is extracting the funded researchers' data and then collecting all the articles that were published by those researchers. This will surely result in an over-estimation of the number of articles. The procedure that we took is based on the assumption that all the grantees should acknowledge the source of funding in the article. The refined data from phases one and two was integrated into a single MySQL[4] table. In the last phase, we used Pajek[5] software to construct the co-authorship networks of the funded researchers and to calculate the network structure variables at the individual level. The calculated network structure indicators were integrated into the database. The final database contains 174,773 records. In the next section, we discuss the methodology.

---

[1] Natural Sciences and Engineering Research Council of Canada, is one of the main funding bodies in Canada. For more information see: http://www.nserc-crsng.gc.ca/index_eng.asp

[2] Scopus is a commercial database of scientific articles that has been launched by Elsevier in 2004. It is now one of the main competitors of Thomson Reuter's Web of Science.

[3] www.scimagojr.com

[4] Open source relational database management system, for more information see: http://www.mysql.com/

[5] Social network analysis software, for more information see: http://vlado.fmf.uni-lj.si/pub/networks/pajek/

Methodology

We first employed social network analysis to construct the collaboration network of the funded researchers and to measure the structural network properties. As the next step, we used statistical analysis to analyze the impact. For this purpose, we considered four different dependent variables that were average team size of the funded researchers (*teamSize*) measured by average number of authors per paper, betweenness centrality (*bc*), and closeness centrality (*cl*). Number of authors per paper has been used in the literature as a proxy for scientific collaboration (e.g. Beaver & Rosen, 1979; Rosenweig et al., 2008). Betweenness Centrality (*bc*) focuses on the role of intermediary individuals in a network. Betweenness centrality of node *k* is measured based on the share of times that a node *i* reaches a node *j* via the shortest path passing from node *k* (Borgatti, 2005). Hence, the more a node lies on the shortest path between two other nodes in a network, the higher betweenness centrality it has that indicates the higher control of the node over other two non-adjacent nodes (Wasserman, 1994). Hence, betweenness centrality of node *k* (bc_k) is defined as follows:

$$bc_k = \sum_{i \neq k \neq j} \frac{\sigma_{ij}(k)}{\sigma_{ij}}$$

where σ_{ij} is the total number of shortest paths from node *i* to *j* and $\sigma_{ij}(k)$ is the number of shortest paths from node *i* to node *j* that contains node *k*. Closeness Centrality (cl) was first proposed by Sabidussi (1966) and is defined based on the shortest path between the nodes in a graph. This measure of centrality considers both direct and indirect connections among the nodes. Hence, the closeness centrality of a node *i* in a graph with N nodes is:

$$cl_i = \frac{1}{\sum_{j \in N-\{i\}} d(i,j)}$$

where $d(i,j)$ is the length of shortest path between the nodes *i* and *j*. Based on the definition, closeness centrality can only be calculated in connected components (graphs) since if the graph is not connected the denominator becomes ∞ and as a result the closeness centrality would be zero which is not informative. To perform the statistical analysis, a regression model was defined for each of the dependent variables and STATA 12[6] data analysis and statistical software was used to estimate the models. The reduced form of the regression models is as follows:

$$\begin{bmatrix} teamSize_i \\ bc_i \\ cl_i \end{bmatrix} = f(avgFund3_{i-1}, avgIf3_{i-1}, avgArt3_{i-1}, avgCit3_{i-1}, dc_i, careerAge_i, d_i)$$

$AvgFund3_{i-1}$ is the average amount of funding that the researcher has received over the past three years. In the literature three-year (e.g. Payne & Siow, 2003) or five year (e.g. Jacob & Lefgren, 2007) time windows have been considered for funding to take effect. We considered both for our models and found that the three-year time window is better suited. As a proxy for the quality of the papers, we added $avgIf3_{i-1}$ to the model that is calculated based on the average impact factor of the journals that the author has published articles in a three year time interval. We also added $avgCit3_{i-1}$ variable to the model that is the average citations for the articles in the past three years as another measure for the quality of the papers. Past productivity of the funded researcher is represented by $noArt_{i-1}$ in the model and is measured as the average number of articles for a researcher in a three year time window. Older researchers in general can be

[6] For more information see: http://www.stata.com/stata12/

more productive (Merton, 1973; Kyvik & Olsen, 2008) due to several reasons like better access to the funding and expertise sources, more established collaboration network, better access to modern equipments. Hence, we included a control variable named $careerAge_i$ representing the time difference between the date of his/her first article in the database and the given year. Degree Centrality (dc) variable was also included in the model which is defined based on the number of ties that a node has (degree) in an undirected graph. Hence, researchers with high degree centrality should be more active since they have higher number of ties (links) to other researchers (Wasserman, 1994). Degree centrality for node i is defined based on the node's degree and then the values are normalized between 0 and 1 to be able to compare centralities:

$$dc_i = \frac{degree\ of\ node\ i}{highest\ degree\ in\ the\ network}$$

In each of the models we used different types of dummy variables. The dummy variable $dInst_i$ represents the type of the affiliation of the funded researcher, whether it is affiliated with academia or non-academia environments. For the Canadian provinces, we defined another dummy variable $dProvince_i$. To compare the impact of different NSERC funding programs $dProg_i$ was also included.

Results

Descriptive analysis

Before turning to the regression models, we first analyze the overall trends of the dependent variables as well as funding, as the main determinant influencing factor of scientific activities (Martin, 2003). As it can be seen in Figure 1-a, average funding received per researcher has followed an increasing trend while after 2003 (vertical line in Figure 1-a) the slope has become steeper indicating a considerable increase in the average amount of funding. In addition, during the first five years of the examined time interval (dashed vertical line in Figure 1-a) we see a steadier trend of the average funding in comparison with the other periods. We will use the vertical lines of average funding in the rest of the figures of this section to see the impact of funding easier. In addition, we define and use *funding periods I*, *II*, and *III* that refer to the funding periods of 1996-2000, 2000-2003, and 2003-2010 respectively.

Researchers publish their results in books or journal articles or present them in scientific conferences to preserve priority for their discoveries and raise their scientific reputation. Although most of the articles were single authored till 1920s (Greene, 2007), today in most of the academic disciplines (except humanities) researchers prefer multi-authorship model due to nature of the big science that requires collaboration and expertise of many individuals (de Solla Price, 1986). Number of authors per paper has been considered as a proxy for scientific collaboration in several studies (e.g. Newman, 2004; Rosenweig et al., 2008). The vertical lines in Figure 1-b show different periods of average funding that was discussed earlier. According to Figure 1-b, it seems that higher funding enables the funded researchers to form larger scientific teams in an aim to increase their productivity. This is quite reasonable since apart from the higher complexity of science the competition among scientists to get access to better resources has also increased, hence the average number of authors per paper is augmenting (Powers, 1988).

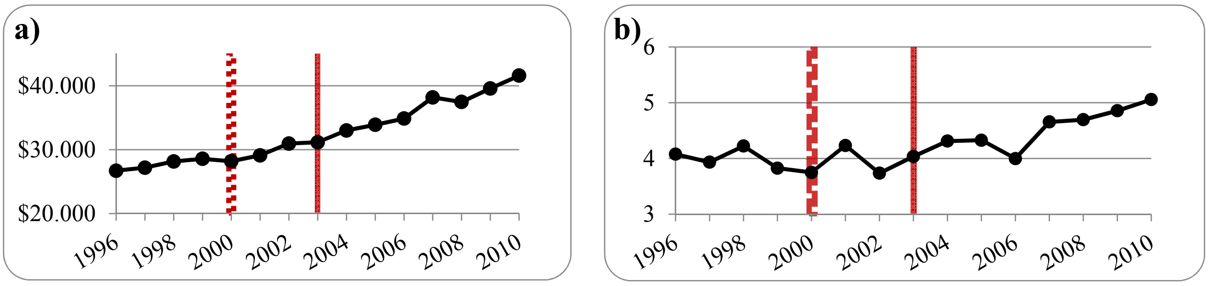


Figure 1. a) Average funding per researcher, b) number of authors per article

The trends of the network structure variables are represented in Figure 2. Except some minor fluctuations, the overall trend of degree centrality is almost steady. However, a significant decline in degree centrality is observed during the years of funding period I. Although the trend of betweenness centrality is steady during the funding period I, it drastically increases within the funding period II maintaining its level in funding III despite some fluctuations. To evaluate the impact of the influencing factors on collaboration more accurately we turn to the regression analysis at the individual level.

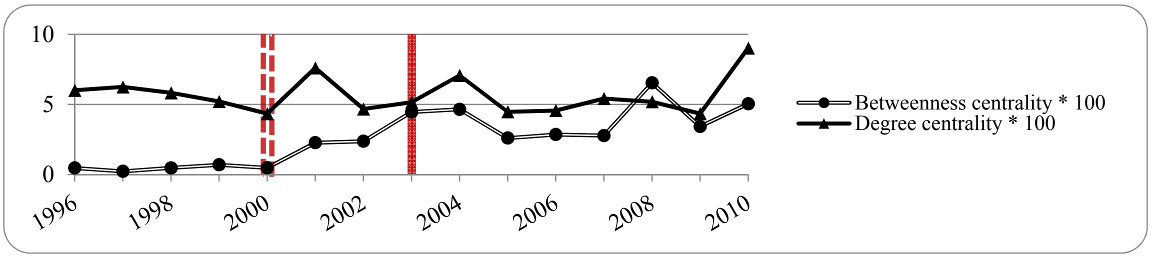


Figure 2. Average betweenness centrality, clustering coefficient, and degree centrality per year

Statistical analysis

In this section, the regression results are presented and discussed for both types of the dependent variables that were discussed earlier.

Average number of authors per paper (teamSize)

The impact of the influencing factors on the scientific team size is analyzed at the individual level. To calculate the team size, we took all the co-authors of a researcher into the account. In all of the regression models, we considered all the combinations of the lags for the variables in the model and used the ones that yield the most robust results. This is similar to the approach of Schilling and Phelps (2007), and Beaudry and Allaoui (2012). We used non-linear time related multiple regressions for the analysis purpose. According to Table 1, the average amount of researcher's funding in the past three years has a significant and relatively high positive impact on overall team size of the researcher. This is in accordance with several studies (e.g. Adams et al., 2005; Gulbrandsen & Smeby, 2005) who found that larger amount of funding will result positively affect the scientific collaboration. As expected, past productivity of the funded researchers (*noArt3*) has also a positive impact on the team size. This may partially highlight the importance of collaboration in scientific activities in a way that highly productive researchers benefit from larger scientific teams. According to the results not only the rate of publications affects the team size, the quality of the works also positively influences the collaboration (*avgCit3* and *avgIf3*). Hence, the results suggest that high quality productive researchers are more collaborative.

Table 1. Regression result, overall team size model

$teamSize_i$	*Coef.*	*Std. Err.*	*t*	*P>|t|*	*[95% Conf. Interval]*	
$ln\_avgFund3_{i-1}$	1.092452^{***}	.2116682	5.16	0.000	.6775817	1.507322
$noArt3_{i-1}$	$.6196625^{***}$	.1215581	5.10	0.000	.3814082	.8579168
$ln\_avgCit3_{i-1}$	1.189371^{***}	.1944058	6.12	0.000	.8083347	1.570407
ln_avgIf3_{i-1}	4.353832^{***}	.3167796	13.74	0.000	3.732943	4.974721
$careerAge_i$	$-.7124596^{***}$	.2184799	-3.26	0.001	-1.140681	-.2842384
$careerAge_i^2$	$.0346901^{***}$	.013419	2.59	0.010	.0083889	.0609913
Affiliations dummy variable						
dAcademia	-8.096576^{***}	1.12634	-7.19	0.000	-10.30421	-5.888946
_cons	.8180767	2.312413	0.35	0.724	-3.71426	5.350413

Notes: * p<0.10, ** p<0.05, *** p<0.01, number of observations: 60,907

We controlled for the career age of the researchers and as expected it was observed that it negatively influences the collaboration. Despite the advantages of collaboration (e.g. better access to resources, internal referring, *etc.*), there are some costs (e.g. finding right partners and research coordination) related to the scientific collaboration (He, Geng, & Campbell-Hunt, 2009). Hence, it seems that as the career age of researchers grow negative impact of costs of collaboration increases in a way that at a certain level senior researchers may tend not to increase their team size. A quadratic term of the career age (*$careerAge^2$*) was added to see the curvature of the relationship and it is seen that the curve of career age is convex (apex at the bottom). To evaluate the impact of the type of the affiliation of the researcher on collaboration, the institution type dummy variable (*dAcademia*) was also to the model that takes value 1 if the funded researcher belongs to the academia environment and 0 if his affiliation is non-academia. As it can be seen, academia funded researchers are significantly different from the non-academia ones and they work in smaller scientific teams in comparison with their non-academic counterparts.

Network structure variables

In this section, using multiple regression analysis at the individual level the impact of influencing factors on betweenness and closeness centralities (dependent variables) is analyzed. Degree centrality of the researchers (*dc*) was added to the model as a proxy of the scientific team size of the researchers. According to Table 2, rate and quality (measured by the average number of citations) of researchers' papers in the past three years have the highest positive impact on their betweenness centrality in the following year. Hence, it can be said that a researcher with more number of publications that are on average of high quality possesses a more central position in the co-authorship network, acting as an influential intermediary in knowledge diffusion and the formation of scientific collaboration. Surprisingly, a negative relation is found between the average impact factor of the journals in which the researchers have published their articles (*avgIf3*) and the betweenness centrality. It seems that the average number of citations is a better proxy for evaluating the quality of the works in the co-authorship network of funded researchers and according to the results not necessarily publishing in higher quality journals may lead the researcher to a more influential position. As expected the average amount of funding received in the past three years (*avgFund3*) has also a positive impact on the centrality of the funded researchers in a way that more funded researchers would be more probable candidates for the central positions of the network. This finding is partially supported by the positive impact of the team size of the researchers measured by their degree centrality (*dc*) since higher amounts of funding may enable researchers to expand their scientific activities that might be resulted in more central positions.

Our findings suggest a negative impact of career age of the researchers on their betweenness centrality indicating that as time passes from the date of the first publication of a researcher, betweenness centrality declines. We also compared the betweenness centrality of the researchers affiliated with academia and non-academia, estimated by the *dAcademia* dummy variable in the model. According to the results, the affiliation of the researchers does not differently affect their central positions and there is no correlation between the type of the affiliation of the researchers and their betweenness centrality. We did the same analysis for the impact of the location of the researchers categorized by different Canadian provinces. We omitted Ontario and defined dummy variables for the remained nine provinces and found that none of the dummy variables of the provinces are significant at the level of 90%. This confirms that locating in one of the other nine provinces does not have a significant different impact from locating in Ontario on the betweenness centrality of the researchers. Finally, we defined dummy variables for the most frequent NSERC funding programs, namely discovery grants, strategic projects, industrial funding, collaborative grant, and tools and equipment grants. The dummy variable of the discovery grants was omitted. It was found that the collaborative grants and strategic projects are significantly and positively different from the omitted dummy variable at the level of 90% and 99% respectively. This partially indicates that researchers who have been funded through collaborative or strategic programs possess in general more central positions in comparison with their counterparts who have been supported by the discovery grants. This finding is completely in line with the definition of the mentioned NSERC funding programs. Specifically for the strategic project grants, the aim is to improve the scientific development in selected high-priority areas that influences Canada's economic and societal position. Hence, these well-defined targeted grants should be allocated to specific reputable researchers, probably with more central positions and higher influential potency.

Table 2. Regression result, betweenness centrality (bc) model

$bc_i * 10^4$	*Coef.*	*Std. Err.*	*t*	*P>|t|*	*[95% Conf. Interval]*	
$ln\_avgFund3_{i-1}$	.4350983***	.0653106	6.66	0.000	.307088	.5631086
$noArt3_{i-1}$	1.409226***	.0332145	42.43	0.000	1.344124	1.474327
$ln\_avgCit3_{i-1}$	.9382845***	.0609348	15.40	0.000	.8188508	1.057718
ln_avgIf3_{i-1}	-.2877661***	.1026326	-2.80	0.005	-.4889287	-.0866036
$dci * 10^4$	.0097551***	.00219	4.45	0.000	.0054626	.0140476
$careerAge_i$	-.0328585**	.0165532	-1.99	0.047	-.0653031	-.0004139
Affiliations dummy variable						
dAcademia	-.0432513	.3988558	-0.11	0.914	-.8250186	.738516
_cons	-5.563956***	.734255	-7.58	0.000	-7.003114	-4.124798

Notes: * p<0.10, ** p<0.05, *** p<0.01, number of observations: 38,974

As the last part of the research, we calculated the closeness centrality of the researchers in the largest component of the co-authorship networks[7] to evaluate the impact of the influencing factors on the closeness centrality (*cl*) of the researchers at the individual level. According to Table 3, average funding (*avgFund3*) positively affects the closeness centrality of the researchers. Hence, it can be said that more funding may enable researchers with high closeness centrality (who are important influencers within their local network) to increase their penetration and prestige. Although a positive affect was observed for the rate of publication (*noArt3*) on the closeness centrality, the relation between the quality of the papers and closeness centrality is not that much clear since the citation based proxy (*avgCit3*) shows a negative impact while the journal impact factor based measure (*avgIf3*) presents a positive effect. It seems that local influencers are not necessarily highly prolific scientists in terms of the quality of their publications. As it was expected, the direct scientific team size of the researchers, measured by degree centrality (*dc*), has a significant positive impact on their closeness centrality since local influencers may benefit from larger team sizes and higher number of connections to empower their penetration within their local community. Based on the results for the *careerAge* and *careerAge*^{2} variables, the impact of the career age on the closeness centrality of the researchers is negative at first. However, approximately after 18 years the overall impact of the career age becomes positive. Therefore, the curve of the career age in the closeness centrality model is convex with the minimum around the age of 18. Hence, it seems that mid-career scientists are more likely to have higher influence within their local community.

Table 3. Regression result, closeness centrality (cl) model

$cl * 10^2_i$	*Coef.*	*Std. Err.*	*t*	*P>|t|*	*[95% Conf.*	*Interval]*
$ln\_avgFund3_{i-1}$	.1694727***	.0269733	6.28	0.000	.1166018	.2223436
$noArt3_{i-1}$	.0215427***	.0034334	6.27	0.000	.0148128	.0282727
$ln\_avgCit3_{i-1}$	-.0734994***	.0263145	-2.79	0.005	-.1250789	-.0219198
ln_avgIf3_{i-1}	.4752498***	.0437594	10.86	0.000	.389476	.5610235
$dc_i * 10^2$	2.593725***	.0546285	47.48	0.000	2.486647	2.700804
$careerAge_i$	-.4191376***	.0261731	-16.01	0.000	-.4704401	-.3678351
$careerAge^2_i$	.0243901***	.0015257	15.99	0.000	.0213995	.0273807
Affiliations dummy variable						
dAcademia	.0596839	.1436248	0.42	0.678	-.2218382	.341206
_cons	8.016092***	.2915756	27.49	0.000	7.444568	8.587615

Notes: * p<0.10, ** p<0.05, *** p<0.01, number of observations: 15,046

Analyzing the results for the dummy variables reveals that academia and non-academic researchers (measured by *dAcademia*) do not have significantly different impact on the closeness centrality. Hence, it is equally likely that local influencers come from industry or academic environments. In addition, researchers who are located in Quebec, British Columbia, Alberta, Saskatchewan, and Manitoba are significantly different from the ones who reside in Ontario. The coefficient is positive for all the mentioned provinces indicating higher closeness centrality of the researchers located in the mentioned provinces in comparison with their

[7] Closeness centrality can be only calculated in the connected networks.

counterparts in Ontario. The coefficient was the highest for the researchers reside in Manitoba. We also compared the impact of different NSERC funding programs for which the discovery grants program was omitted. It was found that the effect is only different for tools and industrial funding programs, with positive and negative coefficients respectively.

Conclusion

In this paper we investigated the impact of funding and other influencing factors like past productivity, team size, and career age of the researchers on their positions and roles within the co-authorship networks. To our knowledge this is the first study that considers the network structure measures as dependent variables and performs the impact analysis on them at the individual level. Analyzing the impact of the influencing factors on the traditional collaboration and scientific team size indicators revealed that funding plays a significant positive role in motivating researchers to collaborate more. This finding is in line with several studies, e.g. Adams et al. (2005) and Gulbrandsen and Smeby (2005). In addition, it was observed that highly productive researchers who are producing high quality papers on average have larger scientific teams. This partially confirms the importance of collaboration in scientific activities in a way that high quality productive researchers tend to be more collaborative. It was observed that the career age of the researchers negatively influences their collaboration that might be due to difficulties in managing the costs of collaboration (e.g. finding right partners and research coordination).

In the second part of the analysis the impact was investigated on the network structure variables. Researchers with high betweenness centrality (gatekeepers) are often critical to scientific collaboration and knowledge diffusion as they can control the flow of information and collaboration. Our result suggest that the past productivity of the researchers in terms of both quantity and quality of the publications along with the average amount of funding available are crucial factors in achieving higher betweenness centrality. Analyzing the impact of degree centrality as a measure of the team size on the betweenness centrality revealed that in the examined co-authorship network higher number of direct connections empowers the role of gatekeepers. Surprisingly, a negative impact was observed for the career age of the researchers on their betweenness centrality that might indicate the considerable role of young gatekeepers in connecting different scientific communities (clusters) and knowledge diffusion in the examined collaboration network.

Researchers with high closeness centrality are identified as important local influencers within their local collaboration network or community. Although they might not be important actors in the entire network, they are highly respected locally as they are on the local short paths of knowledge diffusion. Our results showed a positive impact of funding on the closeness centrality suggesting that local influencers may use more funding to increase their penetration and prestige within their local community. Analyzing the impact of past productivity revealed that local influencers are not necessarily highly prolific scientists specifically in terms of the quality of their publications. However, number of direct connections plays an important role in a way that local influencers can use it to empower their penetration within their local community. Analyzing the impact of the career age showed that the overall impact of the career age becomes positive after 18 years hence it seems that mid-career scientists are more likely to have higher influence within their local community.

We were exposed to some limitations in this paper. Firstly, Scopus and other similar databases are English biased, hence, non-English articles are underrepresented (Okubo, 1997). Another inevitable limitation about the data was the spelling errors and missing values. We measured closeness centrality in the largest component of the co-authorship networks since based on the classic definition of the closeness centrality it can be defined in connected graphs or sub-graphs.

Future works can address this issue by considering the new approaches (e.g. Latora & Marchiori, 2001; Dangalchev, 2006) and comparing the results with the ones of the classic method of calculation of closeness centrality.

References

Adams, J. D., Black, G. C., Clemmons, J. R., & Stephan, P. E. (2005). Scientific teams and institutional collaborations: Evidence from US universities, 1981–1999. *Research Policy, 34*(3), 259-285.

Beaudry, C., & Allaoui, S. (2012). Impact of public and private research funding on scientific production: The case of nanotechnology. *Research Policy, 41*(9), 1589-1606.

Beaver, D. d., & Rosen, R. (1979). Studies in scientific collaboration-part II. scientific co-authorship, research productivity and visibility in the French scientific elite, 1799–1830. *Scientometrics, 1*(2), 133-149.

Borgatti, S. P. (2005). Centrality and network flow. *Social Networks, 27*(1), 55-71.

Bozeman, B., & Corley, E. (2004). Scientists' collaboration strategies: Implications for scientific and technical human capital. *Research Policy, 33*(4), 599-616.

Brad Wray, K. (2006). Scientific authorship in the age of collaborative research. *Studies in History and Philosophy of Science Part A, 37*(3), 505-514.

Dangalchev, C. (2006). Residual closeness in networks. *Physica A: Statistical Mechanics and its Applications, 365*(2), 556-564.

de Solla Price, D. J. (1986). *Little science, big science... and beyond* Columbia University Press New York.

Defazio, D., Lockett, A., & Wright, M. (2009). Funding incentives, collaborative dynamics and scientific productivity: Evidence from the EU framework program. *Research Policy, 38*(2), 293-305.

Godin, B. (2003). The impact of research grants on the productivity and quality of scientific research. *No. 2003. INRS Working Paper.*

Greene, M. (2007). The demise of the lone author. *Nature, 450*(7173), 1165-1165.

Gulbrandsen, M., & Smeby, J. (2005). Industry funding and university professors' research performance. *Research Policy, 34*(6), 932-950.

He, Z., Geng, X., & Campbell-Hunt, C. (2009). Research collaboration and research output: A longitudinal study of 65 biomedical scientists in a New Zealand university. *Research Policy, 38*(2), 306-317.

Heffner, A. G. (1981). Funded research, multiple authorship, and subauthorship collaboration in four disciplines. *Scientometrics, 3*(1), 5-12.

Katz, J. S., & Martin, B. R. (1997). What is research collaboration? *Research Policy, 26*(1), 1-18.

Kyvik, S., & Olsen, T. B. (2008). Does the aging of tenured academic staff affect the research performance of universities? *Scientometrics, 76*(3), 439-455.

Latora, V., & Marchiori, M. (2001). Efficient behavior of small-world networks. *Physical Review Letters, 87*(19), 198701.

Lee, S., & Bozeman, B. (2005). The impact of research collaboration on scientific productivity. *Social Studies of Science, 35*(5), 673-702.

Martin, B. R. (2003). The changing social contract for science and the evolution of the university. *Science and Innovation: Rethinking the Rationales for Funding and Governance.Edward Elgar, Cheltenham,* 7-29.

Merton, R. K. (1973). The sociology of science: Theoretical and empirical investigations. *University of Chicago press*.

Newman, M. E. (2004). Who is the best connected scientist? A study of scientific coauthorship networks. *Complex networks* (pp. 337-370) Springer.

Payne, A. A., & Siow, A. (2003). Does federal research funding increase university research output? *Advances in Economic Analysis & Policy, 3*(1).

Powers, R. D. (1988). Multiple authorship, basic research, and other trends in the emergency medicine literature (1975 to 1986). *The American Journal of Emergency Medicine, 6*(6), 647-650.

Rosenzweig, J. S., Van Deusen, S. K., Okpara, O., Datillo, P. A., Briggs, W. M., & Birkhahn, R. H. (2008). Authorship, collaboration, and predictors of extramural funding in the emergency medicine literature. *The American Journal of Emergency Medicine, 26*(1), 5-9.

Sabidussi, G. (1966). The centrality index of a graph. *Psychometrika, 31*(4), 581-603.

Schilling, M. A., & Phelps, C. C. (2007). Interfirm collaboration networks: The impact of large-scale network structure on firm innovation. *Management Science, 53*(7), 1113-1126.

Sonnenwald, D. H. (2007). Scientific collaboration. *Annual Review of Information Science and Technology, 41*(1), 643-681.

Tijssen, R. J. (2004). Is the commercialisation of scientific research affecting the production of public knowledge?: Global trends in the output of corporate research articles. *Research Policy, 33*(5), 709-733.

Tijssen, R. J., van Leeuwen, T. N., & Korevaar, J. C. (1996). Scientific publication activity of industry in the netherlands. *Research Evaluation, 6*(2), 105-119.

Ubfal, D., & Maffioli, A. (2011). The impact of funding on research collaboration: Evidence from a developing country. *Research Policy, 40*(9), 1269-1279.

Wasserman, S. (1994). *Social network analysis: Methods and applications* Cambridge university press.

Benefits of scientific collaboration

Barbara S. Lancho Barrantes

Faculty of Information Science and Communication.
University of Extremadura. Plazuela Ibn Marwan, 06071 Badajoz (Spain)
bslancho@unex.es

Abstract

Scientific papers that include international collaboration obtain a greater impact and may gain in citations in general because they have multiple “immediate surroundings,” and of course, because of their greater quality or prestige. Undoubtedly there is a greater impact on the authors "immediate surroundings", but this does not necessarily have to coincide with geographical proximity or with their national environments, which fade in importance as the collaborative environment expands. In short, one can say that science knows no frontiers.

On the other hand, collaborating with a country increments the citation received from it. But some collaborating countries provide large increases in references in this sense than others, and likewise some countries receive greater increments of citation from their partner countries than others. This oral presentation also emphasizes the origin of the citation obtained by different countries that work together on a paper (collaborating) and the destination of the references realized by these countries. Taking into account also the production of a country therefore the higher the production of one country. Moreover the increase of citation that countries provided to their collaborating countries results different among areas (Medicine, Social Sciences, Engineering and Physics).

Keywords: Citation analysis, Citation increment, Scientific collaboration, Scientific collaboration in subject areas, Scientometrics

Introduction

In the scientific world it has been said that great levels of collaboration lead to high levels of impact, greater quality of the papers published, and greater productivity of the authors in their personal research areas (Narin, Stevens, & Whitlow, 1991; Glänzel, 2001; Leimu & Koricheva, 2005; Katz & Hicks, 1997; Persson, Glanzel, & Danell, 2004; Hsu & Huang, 2010) The effect of collaboration on scientific impact appears to be more positive in the“hard”sciences such as physics and astronomy, than in the “soft” sciences such as humanities or social sciences (Bandyopadhyay, 2001; Moed, Bruin, Nederhof, & Tijssen, 1991; Bridgstock, 1991).

Therefore where does that impact come from? where references are intended? Does it come mainly from the countries included in the collaboration? Could the citation be because of the international partners in the collaboration?

This increment in citation can depend on the type of scientific collaboration. For example, the greatest increment in citation comes from collaboration with institutions of different countries (Narin et al. 1991; Katz and Hicks 1997; Goldfinch et al. 2003).

The main hypothesis of the present work is that science knows no frontiers—that there is no national citation bias.

Data and Methods

Scopus is used as the data source for the computation of the indicators because it best represents the overall structure of world science at a global scale. It covers most of journals included in Thomson Reuters ScientificWeb of Science (WoS) and more and it coverage is statistically balanced in terms of subjects, countries, languages, and publishers.

Results

In this section is included only the discussion of the results to answer some of the issues. The full results have been exposed and developed in detail in each publication (Lancho et al. 2013a; Lancho, Guerrero and Moya, 2013b) along with the sources, methodologies, constraints encountered.

Table 1. The 20 countries ranked by total number of papers, with the corresponding percentage of papers in collaboration, percentage of citations from collaborating countries, and percentage of references to collaborating countries.

Country	Papers 2004	Papers 2005 to 2007	% Collab. Papers 2004	% Collab. Papers 2005 to 2007	% Cit. from Collaborators	% Ref. to Collaborators
United States	410521	1359565	23.28	25.29	15.57	14.36
China	113292	552902	16.09	13.99	24.18	32.65
Japan	110622	347570	19.79	20.94	35.30	34.89
United Kingdom	107143	362408	37.46	38.99	30.86	30.24
Germany	98949	326678	39.65	41.47	30.65	29.49
France	69591	232091	41.64	43.33	32.51	31.07
Canada	55929	196960	39.27	40.78	39.06	40.04
Italy	54112	183950	34.81	36.64	33.39	33.67
Spain	39776	144564	32.88	34.89	33.29	32.92
Australia	35886	127554	38.03	38.85	36.18	35.26
Russian Fed.	35000	100774	30.93	33.20	34.89	32.77
India	33169	128505	18.22	18.28	33.61	33.08
Republic Of Korea	31332	120173	25.00	25.45	38.68	42.53
Netherlands	30168	104253	45.14	45.90	35.05	35.36
Switzerland	22049	75615	55.05	56.95	35.58	35.59
Brazil	21658	88335	27.75	25.35	36.16	35.50
Sweden	21280	69274	45.75	48.90	34.38	32.30
Taiwan	21071	81275	17.49	18.53	38.25	42.70
Poland	20378	67225	31.32	31.18	36.80	34.40
Turkey	18170	65901	16.75	15.39	32.61	36.60

In Table 1 one can observe that the United States gets only a small percentage of citations from its collaborators and in turn provides them with only a small percentage of references. We understand this to be because of its large production and hence the large number of national references involved, i.e., because U.S. production is so large, its domestic citations reduce the percentage from other countries (and analogously with regard to references).

However China, despite having only a small percentage of collaboration (which even decreased in the second period), devotes a major proportion of its references to collaborating countries.

The countries with the highest percentages of citations received from their collaborating countries and the highest percentages of references given to those countries are Canada, Republic of Korea, Australia, and Taiwan. These are not countries characterized by a great volume of scientific production.

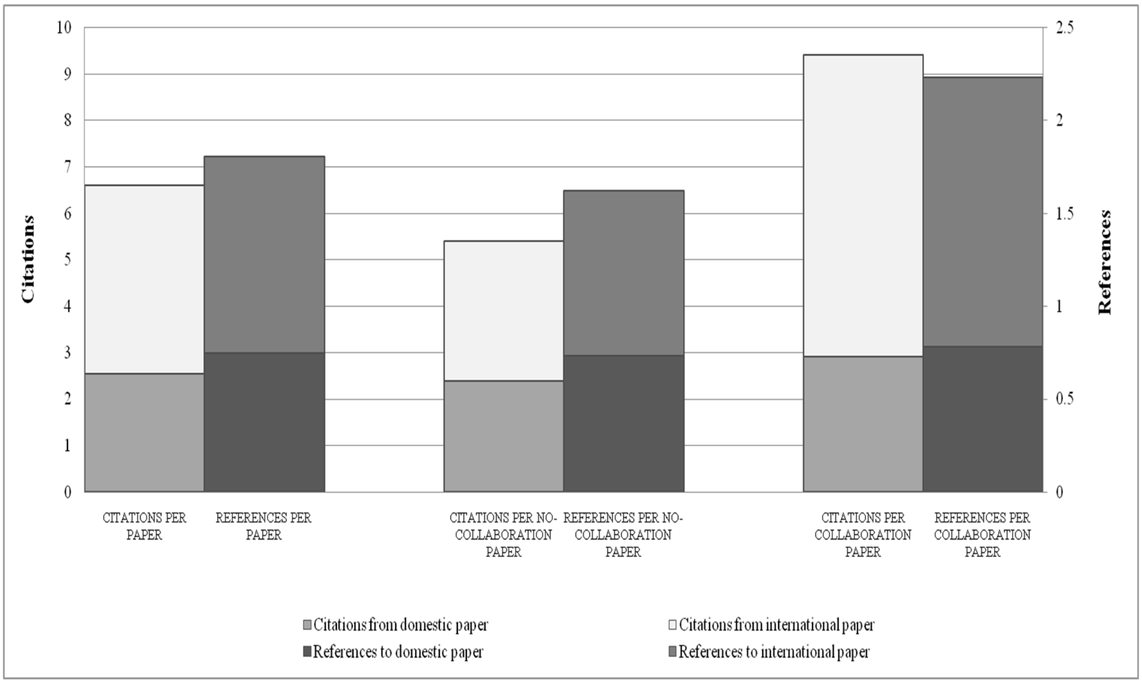


Fig. 1. Comparative relation of weighted average citations (per paper overall, per paper without collaboration, and per paper with collaboration) and references (per paper overall, per paper without collaboration, and per paper with collaboration) of the 20 countries more productive in 2004, distinguishing in both cases between domestic and nondomestic journal articles.

The value of citations per collaboration paper is considerably greater than those of the citations per no-collaboration paper and citations per paper overall, with this difference originating mainly from nondomestic papers. The case is similar for the indicators relating to references, although the differences are less marked. The average of non- domestic references is far greater than that of the domestic references in all cases, and the average references per collaboration paper is considerably greater than those of the references per no-collaboration paper and per paper overall.

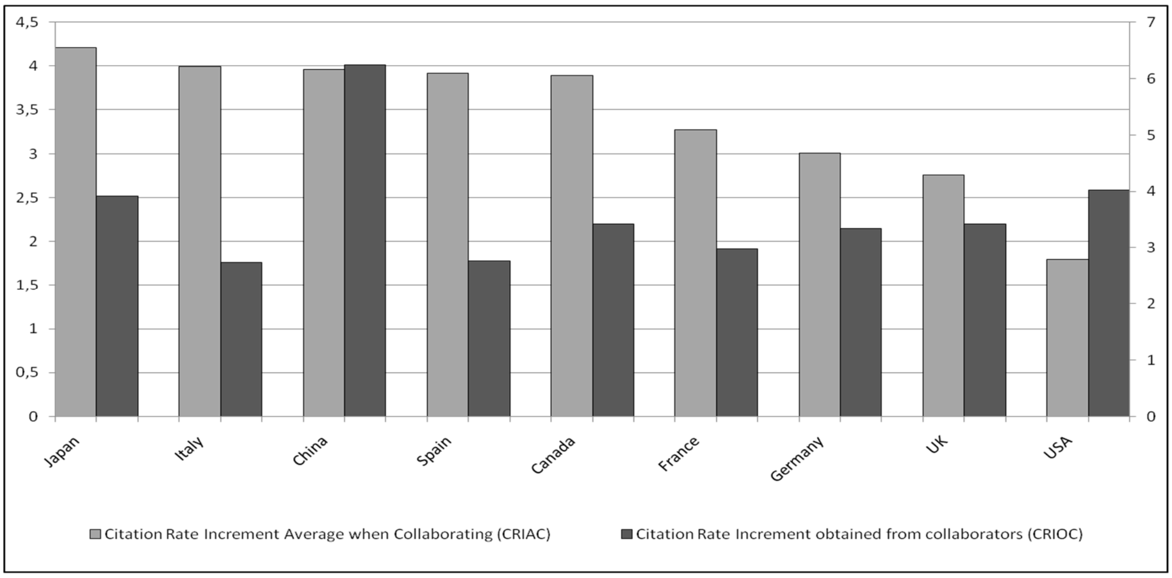


Fig. 2. Comparison between the CRIAC and the Citation Rate Increment Obtained from Collaborators (CRIOC) of the nine countries with the greatest production in 2004. Ordered by CRIAC

When a country collaborates, it obtains a CRIC. In every case studied, a greater percentage of citations was received from a country with which it collaborated than in the case of non-collaboration. But some collaborating countries presented higher values of the CRIAC (to their collaborator countries) than others.

Coincidentally the USA, China, and Japan, the countries with the greatest production in 2004, were the countries with the greatest increment in the number of citations obtained from the countries, but not equally from all the countries with which they collaborated.

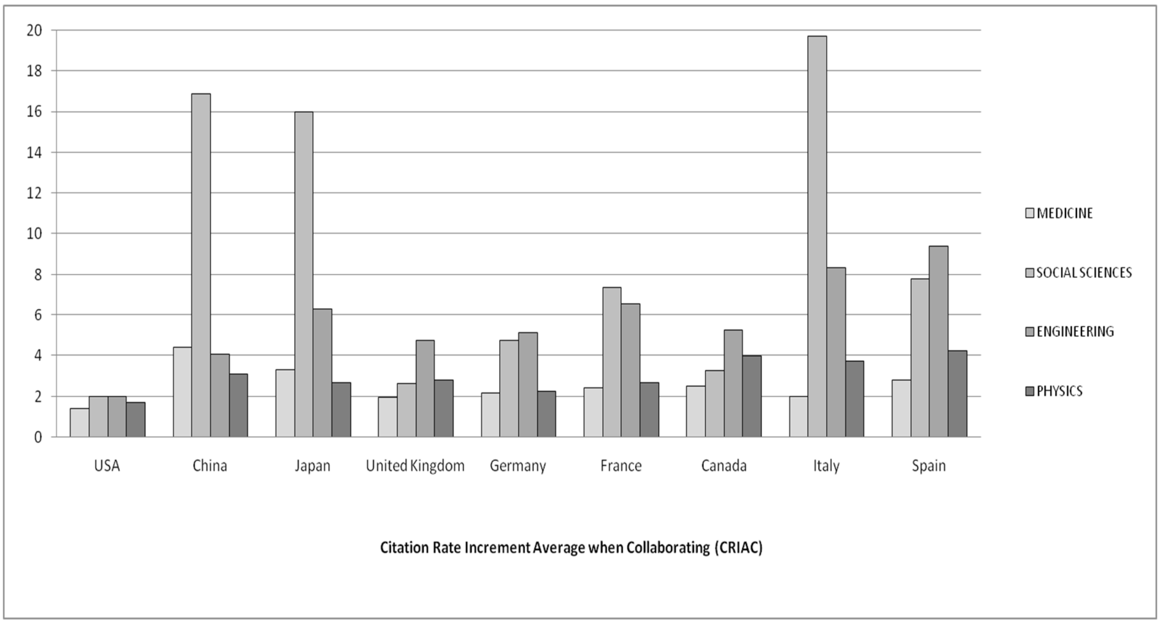


Fig. 3.Citation Rate Increment Average when Collaborating (CRIAC) of the 9 countries with the greatest production in 2004 in the Medicine, Social Sciences, Engineering, and Physics specific areas.

However, some countries had a greater Citation Rate Increment Average when Collaborating (CRIAC) towards their collaborating countries in some areas than in the general case. Certainly, in Social Sciences, Italy, China, and Japan. In Engineering, Spain, Italy, and France almost doubled the increment of the general case. In the areas of Physics and Medicine, the countries deviated little from the general case – their citation rates to a country increased some fourfold when collaborating.

In the case of the USA, there were no differences between the general case and the different specific disciplines. In both cases, the citation rate to a country increased by about 2 times (on average) with collaboration. It was not only the country with the highest total impact in general. but also in all four specific disciplines studied except Social Sciences, in which it ranked second after the UK.

In the different scientific disciplines, there were collaborating countries that received a greater Citation Rate Increment Obtained from Collaborators (CRIOC) than others.

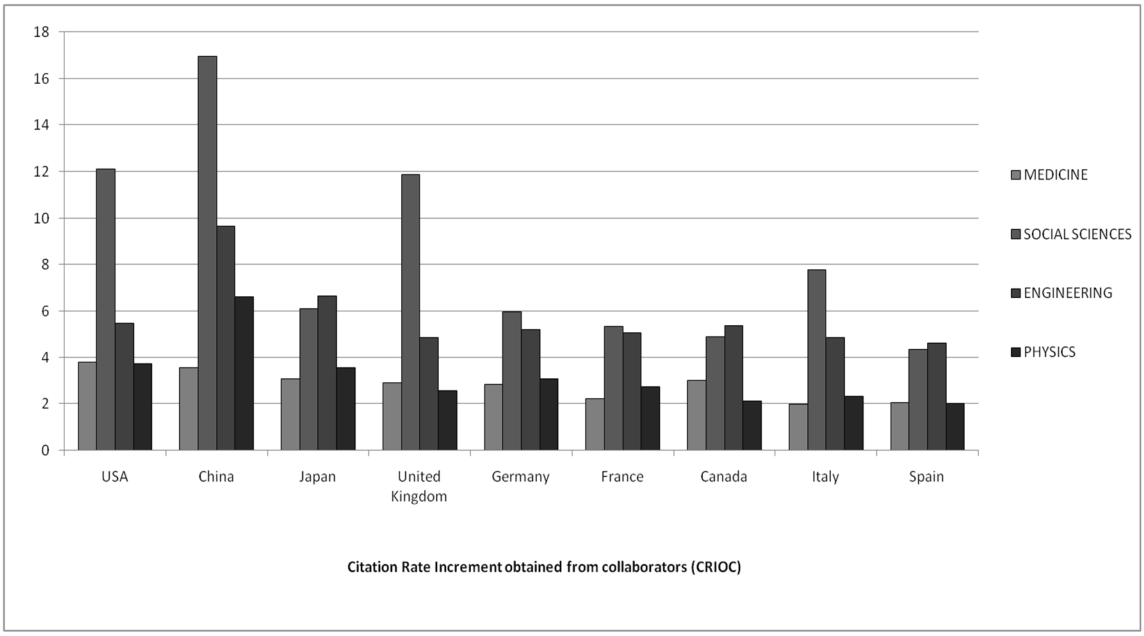


Fig. 4. Citation Rate Increment Obtained from Collaborators (CRIOC) of the 9 countries with the greatest production in 2004 in the Medicine, Social Sciences, Engineering, and Physics specific areas.

In Social Sciences, China, the USA, and the UK received an increment from their collaborators that almost tripled that found in the general study. In Engineering, the values for China and Japan almost doubled those of the general study, perhaps partially because they had a lower impact in this discipline than overall. The areas of Medicine and Physics were those with the lowest values of CRIOC, even lower than in the general study. The only special case was China in Physics, with a value that surpassed that of the general study. Surely for China the impact in Physics exceeded that of the general studies, while the contrary was the case for the other three scientific areas considered.

Interpretation

Scientific collaboration, particularly international scientific collaboration, leads to high levels of impact and large number of citations.

The number of citations per collaboration paper is significantly greater than those of the citations per no-collaboration paper and citations per paper in general, with this difference originating mainly from nondomestic papers. The case is similar for the indicators relating to references, although the differences are less marked.

When a country is involved in collaborations, it receives a positive Citation Rate Increment of the Colaborator (CRIC). In all cases, a greater citation increment was received from collaborating countries than from non-collaborating countries. But some countries had higher values of the CRIAC (to their collaborating countries) than others, and there were also countries receiving a greater Citation Rate Increment Obtained from Collaborators (CRIOC) than others. There seemed to be a tendency for the countries with greater impact to have a smaller increment from collaborating.

In the four scientific disciplines selected for in-depth study, in all there was a positive CRIC. But in Social Sciences, some countries had a CRIAC which was greater than that of the general case by a factor of three, and in Engineering by a factor of nearly two. Indeed, the areas of Social Sciences and Engineering were those in which collaborating countries received greater Citation Rate Increments Obtained from their Collaborators (CRIOC), while the areas of

Physics and Medicine received lower CRIOC increments, even less than in the general studies. There seemed to be a trend in these countries for the increment to be less for collaborations in the disciplines of greater impact.

References

Bandyopadhyay, A.K. (2001). Authorship pattern in different disciplines. *Annals of Library and Information Studies*, 48(4), 139–147.

Bridgstock, M. (1991). The quality of single and multiple authored papers– An unresolved problem. *Scientometrics*, 21, 37–48.

Glänzel, W. (2001). Domestic characteristics in international scientific coauthorship relations. *Scientometrics*, 51(1), 69–115.

Hsu, J.W., & Huang, D.W. (2010). Correlation between impact and collaboration. *Scientometrics*, 86(2), 317–324.

Katz, J., & Hicks, D. (1997). How much is a collaboration worth? A calibrated bibliometric model. *Scientometrics*, 40(3), 541–554.

Lancho-Barrantes, B. S., Guerrero-Bote, V. P., Chinchilla-Rodríguez, Z., & Moya-Anegón, F. (2013a). Citation flows in the zones of influence of scientific collaborations. *Journal of the American Society for Information Science and Technology*, 63(3), 481–489.

Lancho-Barrantes, B. S., Guerrero-Bote, V. P. & Moya-Anegón, F. (2013b) Citation increments between collaborating countries. *Scientometrics*. 94 - 3, pp. 817 - 831.

Lancho, B., Guerrero, V., & Moya, F. (2010a). The iceberg hypothesis revisited. *Scientometrics*, 85(2), 443–461.

Lancho, B., Guerrero,V., &Moya F. (2010b). What lies behind the averages and significance of citation indicators in different disciplines? *Journal of Information Science*, 36(3), 371–382.

Leimu, R.,&Koricheva, J. (2005). Does scientific collaboration increase the impact of ecological articles? *Bio Science*, 55, 438–443.

Moed, H., Bruin, R., Nederhof, A., & Tijssen, R. (1991). International scientific co-operation and awareness within the European Community: Problems and perspectives. *Scientometrics*, 21(3), 291–311.

Narin, F., Stevens, K.,&Whitlow, E. (1991). Scientific cooperation in Europe and the citation of multidomestically authored papers. *Scientometrics*, 21(3), 313–323.

Persson, O., Glanzel,W., & Danell, R. (2004). Inflationary bibliometric values: The role of scientific collaboration and the need for relative indicators in evaluative studies. *Scientometrics*, 60(3), 421–432.

The Application of Funding Acknowledgment on the Path Analysis of Knowledge Dissemination of Granted Researches

Carey Ming-Li Chen

Science & Technology Policy Research and Information Center,
National Applied Research Laboratories
14F, No.106, Sec.2, Heping E Rd., Taipei 10636, Taiwan
carey.mlchen@gmail.com

Abstract

For funding agencies and relevant stakeholders, not only how to identify journal articles are the outputs of granted researches to trace back the performance is a critical issue, but also how to evaluate whether granted researches make their impact to create social accountability is widely discussed by the whole society. Hence, this study tries to undertake a preliminary study by utilizing bibliometric analysis and multivariate statistical analysis to solve these problems. Meanwhile, funding acknowledgement and citation between journal articles and patents are applied to find the possible pathway of knowledge dissemination of granted researches. As a result, a total of 24,248 granted research which funded by National Science Council in Taiwan is recognized, and up to almost 90% of them have already made scientific impact, and 83 granted researches even have influenced on technological inventions. The characteristics of these granted researches with different pathways of knowledge dissemination are significant different. This study provides more understandings on the relationship between characteristics of granted researches and their pathway of "making impact", and hope the result brings implication to policy decision-making process.

Introduction

How to trace back the performance of granted research projects or identify the research sponsorship is always a critical issue for funding agencies and relevant stakeholders (Wang & Shapira, 2011). Especially for funding agencies, it would be very helpful to get the information of the ratio of output to input via identifying which publications or patents are the outputs of the granted researches. The analysis result would help funding agencies to understand the performance of granted researches better and help to assess whether the initial planning of resource allocation bring in the optimal effectiveness or not. Besides, since the concept of social accountability have brought out, more and more researchers are asked to think the possibility of making contribution on well-being of the whole society, particularly the researches which were funded by public sector are required to generate or demonstrate their 'social impact' (Watson, 2010). Therefore, how to keep tracing the process of knowledge diffusion has become more important eventually. To understand the process of knowledge flow, tracing back and identifying the outputs of granted researches and following up how these results make their impact on industry is the objective of this study, and the result may provide better understanding on the path that research results have disseminated. Moreover, the implication of this study will help shape research policy and funding policy in the future.

Research framework and method

To understand how knowledge disseminates and the process of knowledge industrialization, this study decides to analyze the path of knowledge dissemination of granted researches via two-stage-analysis. The first stage of this study is to identify the outputs of granted researches. Since more and more funding agencies have started to ask researchers who received research grants to declare the detail of financial support from which funding agency and mention it in the section of acknowledgement in journal articles when they published the research results (Giles & Council, 2004) recently, funding acknowledgement is considered to be the approach to trace back the performance of granted researches. In fact, funding acknowledgement is the statement in journal article that authors try to express their gratitude about financial support they get during the research process. The statement usually contains two parts, the name of funding agency and the grant number. From the perspective of bibliometrics, funding acknowledgement is the ways to observe the process that how researchers create and construct the knowledge. Moreover, it is viewed to be the method to trace granted researches' outputs in terms of publications because of the name of funding agencies and other information of financial support are included in the content of funding acknowledgement (Costas & van Leeuwen, 2012). Hence, this study tries to analyze the content and extract the grant number to help to verify the sponsorship to get more information about grants.

The second stage of this study is to continue tracing how the results of granted researches make their impact on other academic researches and the industry, and the concept of linkage between science and technology is attempted to apply. Based on the result of Mansfield's study in 1991, a certain number of companies thought some industrial innovation and process innovation could not be done in time without the contribution that were made by the scientific researches. Under this circumstance, the scientific papers and patents are considered as important variables to examine knowledge flow between science and technology, and utilize the citation between journal articles and patents, like non-patent reference, as proxy to observe how academic researches make an impact on industry (Schmoch, 1997). Non-patent reference is one of citation type and exists in many forms, like journal articles, technical reports, or other publish documents (Callaert, Van Looy, Verbeek, Debackere, and Thijs, 2006), and it is added by inventors mostly. In fact, the number of non-patent reference cited in patents has increased rapidly in recent years, and this result indicates that the linkage between science and technology, or even the relationship between academics and industry may be connected more closely (Lee, Chen & Su, 2013). Therefore, this study decides to apply non-patent reference as footprint of knowledge flow to help understand the path of knowledge dissemination from granted researches to technology development.

The research framework is illustrated in Fig.1. To create the linkage between granted researches and their outputs, funding acknowledgement is utilized here to help identify the path. After the name of funding agency and the grant number revealed from journal articles are extracted, the process of data mapping is operated to link the grant information and bibliographic information together. Then the dataset of journal articles is linked to the database of United States Patent and Trademark Office (USPTO), to identify which journal articles are cited by patents that issued by U.S.

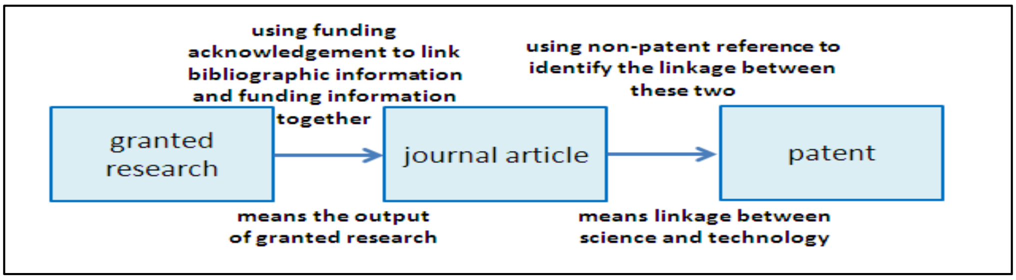


Figure 1. Research framework

Research results

The characteristics of granted researches

This study tries to utilize funding acknowledgement to trace back how granted researches disseminate their results and identify the citation relationship between journal articles and patents to know whether the results of the granted researches are applied in technology development. After the process of data analysis, a total of 39,630 journal articles with funding acknowledgement were identified at first, and these journal articles were the output from a total of 24,248 granted researches which were mainly granted by National Science Council in Taiwan. Besides, these 39,630 journal articles were cited by 92 patents which were issued by USPTO. To get more information, this study tries to find out which granted researches have influenced to the phase of technology development as well.

The characteristics of granted researches are examined, and the result is listed in Table 1. The most of granted researches were basic research, the proportion is 49.04%, and the following is applied research (40.03%). The granted researches which R&D activities were technology development and commercialization are quite few, and the reason might be related to the fact that the mission of National Science Council is to mainly fund pure academic researches instead of researches related to industrial applications. The category of funding department is analyzed in this section as well. The researches which were granted by department of engineering and technology have highest proportion, it is 45.49%, and the followings are department of life science (30.50%) and department of natural science (17.11%). In addition to this, these granted researches produced 2.23 journal articles on average.

Table 1. The characteristics of granted researches

Variable	(n=24,248)
Research type (%)	
Basic research	49.04
Applied research	40.03
Technology development	7.99
Commercialization	2.94
Funding department (%)	
Life science	30.50
Engineering and technology	45.49
Human & Social Science	4.03
Nature science	17.11
Science education & international collaboration	1.29
Other	1.58
Average of journal articles	2.23

The pathways on knowledge dissemination of granted researches

To find out how granted researches disseminate the knowledge and make their impact, this study tries to consider the occurrence of citation as the key elements to observe and generalize the possible pathways. Since the main outputs of granted researches are journal articles, than the citation between other journal articles, even citation between journal articles and patents, have eventually become one important proxy to examine whether the granted researches make their influence on later academic researches or industrial applications. In other words, this study tries to apply whether the granted researches' outputs have been cited or not to be the criteria to classify the granted researches. Hence, once specific granted researches' outputs have not been cited yet, then these granted researches are classified into the group which called "not making impact yet." If the journal articles' citation frequencies are larger than 0, which means the research finding has disseminated to other researches, then this kind of granted researches is named as "making scientific impact." Moreover, if the journal articles have been cited by patents, in terms of non-patent reference in patent inventions, then the granted researches are considered as having influence on industry and therefore named as "making technological impact." The possible pathways of how granted researches disseminate their knowledge are drawn in Fig. 2.

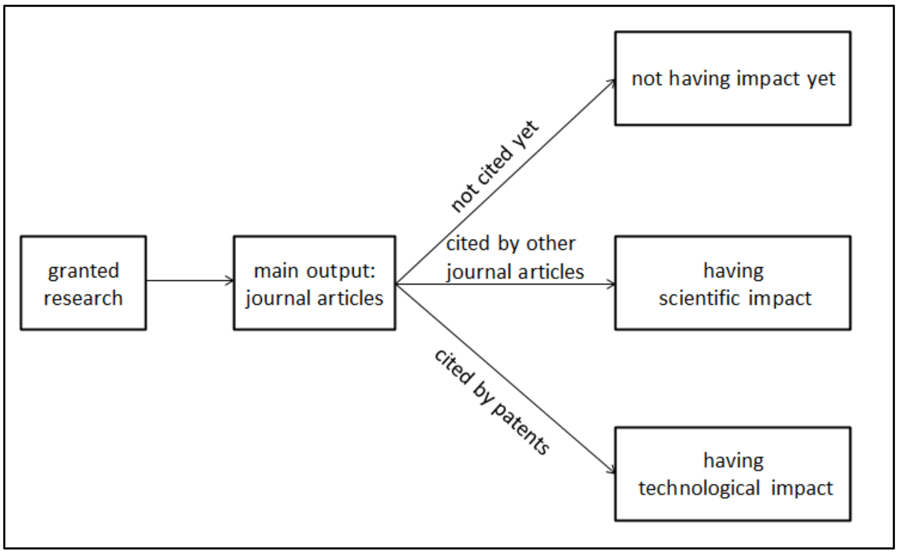


Figure 2. The pathways of knowledge dissemination of granted researches

To identify each granted research's pathway of knowledge dissemination, the descriptive statistics is utilized. The obtained result is listed in Table 2. A total of 2,570 granted researches are classified into Group 1 due to their main outputs, journal articles, which have been not cited by other journal articles yet. The share of it is 10.60%. This group of granted researches produced 1.12 journal articles on average. Group 2 is "having scientific impact", a total of 21,595 granted researches belong to this group, and they produced 2.35 journal articles on average. The result indicates that up to 89.06% of granted researches had journal articles which have been already cited by other journal articles, and made their scientific impact. Group 3 is the granted researches which had journal articles and they have been cited by patents which issued by USPTO. A total of 83 granted researches are classified in this group. This result indicates that only 0.34% of granted researches have made influence on technological inventions, more precisely speaking, they have even made impact on the process of invention, which became the patents which were issued by USPTO respectively. That means the research findings of these granted researches were published in journal articles by the researchers at first, and then the knowledge have influenced on other researches, so they were chosen to be cited, furthermore, these journal articles were cited by technological inventions, in terms of non-patent references in patents, thus it means these findings have made the impact on industrial application as well.

Table 2. Three groups based on pathway of knowledge dissemination

	Group 1	Group 2	Group 3
	"not having impact yet"	"having scientific impact"	"having technological impact"
No. of granted researches	2,570	21,595	83
Shares (%)	10.60	89.06	0.34
AVG. journal articles	1.12	2.35	4.42

The difference on characteristics among there groups of the granted researches

To examine the characteristics of these three groups of granted researches, statistical analysis is applied here. The result is demonstrated in Table 3. The main research type of Group 1 was applied researches, it is 47.32%. Table 3 also indicates that over half of granted researches from Group 2 were basic research, and 39.17% of them were applied research. About Group 3, among of these 83 granted researches, 51.81% were basic research, 38.55% were applied research, and only 9.64% were technology development. This result shows that basic research surprisingly is the type that has disseminated research findings and knowledge to the process of patent invention. Besides, 61.45% of granted researches in Group 3 were financial supported by department of engineering and technology, 24.10% of them were granted by department of life science and rest of them belonged to department of natural science and other department. However, this result is needed to do cross analysis on research type and funding department among these three groups later in order to verify whether different research type or funding department have made differences on the path of knowledge dissemination.

To find out whether these granted researches with different pathways of knowledge dissemination are the same on the characteristic or not, analysis of variance (ANOVA) is utilized to examine the significant difference. Based on the result listed in Table 3, there exists significant difference in research type (χ^2=162.94, p-value < 0.0001), however, the result surprisingly indicates that the type of granted research in Group 1 was mainly applied research, and Group 2 and Group 3 belong to basic researches. About funding department, although the researches among three groups were granted by department of engineering and technology mainly, the proportion is still different significantly (χ^2=446.68, p-value < 0.001). Moreover, the average of number of journal articles which were produced by the granted researches is significant different as well (F=278.21, p-value < 0.001). Therefore, it is concluded that granted researches with different pathway of knowledge dissemination might be related to their characteristics.

Table 3 The characteristics of three groups of granted researches

	Group 1 not having impact yet	Group 2 having scientific impact	Group 3 having technological impact	Statistic test
Research type (%)				χ^2=162.94***
Basic research	37.90	50.35	51.81	
Applied research	47.32	39.17	38.55	
Technology development	11.67	7.54	9.64	
Commercialization	3.11	2.94	0.00	
Funding department (%)				χ^2=446.68***
Life science	14.40	32.44	24.10	
Engineering and technology	61.75	43.50	61.45	
Human & Social Science	5.37	3.88	0.00	
Nature science	15.18	17.35	13.25	
Science education & international collaboration	1.67	1.25	0.00	
Other	1.63	1.58	1.20	
AVG. journal articles	1.12	2.35	4.42	F=278.81***

Note: *** indicates p-value <0.001

Conclusions

This study tries to utilize funding acknowledgement to trace back the performance of granted researches which were mainly funded by National Science Council in Taiwan. Since the concept that granted researches should take social accountability by disseminating their research findings or knowledge to others has widely discussed, this study tries to find out the possible pathways of knowledge dissemination of granted researches. Three pathways are identified based on the frequency of citation between journal articles and patents. There are named as "not having impact yet", "having scientific impact", and "having technological impact."

Up to almost 90% of granted researches have already made impact on other researches because their outputs in terms of publications were cited by other journal articles, but only 0.34% of granted researches have made technological impact on the \patent inventions. The result also surprisingly shows that basic research is the main type which has disseminated their knowledge to the process of invention; only 9.64% of granted researches in Group 3 in terms of "having technological impact" were technology development. At the meantime, all characteristics are different significantly existing among three groups, it is therefore concluded that granted researches with different pathways of knowledge dissemination might be related to their characteristics.

Research limitations

However, there still are several research limitations. Since the dataset in this study is granted researches which funded by Taiwanese funding agencies, they might imply the outputs of the researches were published in domestic journals, moreover, cited by patents which were issued by Taiwanese patent offices, so the proportion of granted researches with different pathways of knowledge dissemination might be underestimated due to the dataset have not contained these citation relationships among domestic journal articles and patents. In addition, whether the citation relationship between journal articles and patents is able to tell the whole story of knowledge dissemination should be more cautious, it might need more evidences to be able to describe the real process of knowledge dissemination of granted researches more completely, thus this study suggests that further studies apply qualitative analysis to fill this research gap. As a result, it is therefore able to provide more precise policy implication to funding agencies when making decision of resource allocation.

Acknowledgement

The author would like to thank Yen-Ting Lin for the help of data cleaning process.

References

Callaert, J., Van Looy, B., Verbeek, A., Debackere, K. & Thijs, B. (2006). Trace of prior art: An analysis of non-patent references found in patent documents. *Scientometrics, 69*(1), 3-20.

Costas, R. & van Leeuwen, T.N. (2012). Approaching the "reward triangle": General analysis of the presence of funding acknowledgements and "peer interactive communication" in scientific publications. *Journal of the American Society for Information Science and Technology, 63*(8), 1647-1661.

Giles, C.L. & Councill, I.G. (2004). Who gets acknowledged: Measuring scientific contributions through automatic acknowledgement indexing. *PNAS, 101*(51), 17599-17604.

Lee, P-C., Chen, C.M-L. & Su, H-N. (2013). Evaluating Application of Non-Patent Reference. *Journal of Technology Management, 18*(3), 1-32. (in Chinese)

Mansfield, E. (1991). Academic research and industrial innovation. *Research Policy, 20*(1), 1-20.

Schmoch, U. (1997). Indicators and the relations between science and technology. *Scientometrics, 38*(1), 103-116.

Wang, J. & Shapira, P. (2011). Funding acknowledgement analysis: An enhanced tool to investigate research sponsorship impacts: The case of nanotechnology. *Scientometrics, 87*(3), 563-586.

Watson, C. (2010). Accountability, transparency, redundancy: Academic identities in an era of ‘excellence’. *British Educational Research Journal, 37*(6), 955-9.

Excellence, merit and research team size: a library and information science case study

Carlos Olmeda-Gómez*, María Antonia Ovalle-Perandones*, Juan Gorraiz** and Christian Gumpenberger**

*Carlos III University of Madrid
carlos.olmeda@uc3m.es, antonia.ovalle@uc3m.es

**University of Vienna
juan.gorraiz@univie.ac.at, christian.gumpenberger@univie.ac.at

Abstract

This paper provides a comparative international overview of partnering in library and information science research, a social sciences subject category whose dynamics and focus are of particular interest for this community. This multinational comparative study embraces 10 years of data and covers different types of partnering at several levels of aggregation and output tiers. It aims to shed more light not only on the issue of the possible advantage of co-authorship for the intents and purposes of citation, but also on the intensity and development of collaboration over the last 10 years.

Introduction

The increase in co-authorship is well documented at all levels of analysis and reflected in the analyses conducted (Glänzel, 2001, 2002; Persson, Glänzel, and Danell, 2004; Wutchy, Jones and Uzzi, 2007).

One controversial factor from the bibliometric standpoint is the study of the possible relationship, in a given discipline, between the merit or impact of the papers published, measured as the number of citations received after publication, and the number of co-authors. The debate stems from the ambiguity of the findings published, according to which no direct relationship can be drawn between team size and impact (Bridgstock, 1991; Rousseau, 2001; Franceschet and Costantini, 2010; Craig Finlay, Ni, and Sugimoto, 2012).

Objectives

This study analyses degrees of co-authorship (i.e., number of authors) in a given subject area and the citations received at each level, for both total output and what may be regarded as the "excellence" tier. The analysis, based on a case study in the area of library and information science (LIS), purports to answer the following questions:

- What degree of co-authorship can be found in worldwide LIS output, by countries and universities, both as a whole and in the "excellence" tier?
- Can significant differences in impact be identified by degree of collaboration in the total or "excellence" tiers?

Methodology

The present study analysed output in the Scopus subject category "Library and Information Science" (LIS) compiled from the Scimago Institutions Ranking (http://www.scimagoir.com) for the years 2003-2012. Two levels of analysis were considered: countries and universities. The analysis addressed two types of output. First, the total output (i.e., regardless of the number of authors, labelled total) in this category attributed to the 30 most productive countries and the 20 universities with the most influential output were calculated from the entire world output in

2003-20012. The issue was also researched at a narrower level, covering the "excellence" or 10 % most frequently cited papers (labelled as top 10 %) published in journals listed under this Scopus subject category, irrespective of the number of authors (Adams, Gurney and Marshall, 2007; Gorraiz, Reimann and Gumpenberger, 2012; Bornmann, Moya-Anegón and Leydesdorff, 2012). This tier contains widely cited papers and the universities with the largest output in this category, which are the ones with the highest impact indicator for their research (Waltman, et al., 2012). This approach circumvents inherently biased citation distributions (Seglen, 1992; Albarrán et. al, 2011), since all the papers in this second group would by definition lie in the same tier.

The full counting method was used to attribute Scopus-listed papers to countries or institutions, based on the authors' institutional affiliations. In other words, if an institution or country appeared in the affiliation field, it was attributed to that country or institution (where there were more than one, all were equally weighted as 1). Using the world data, international collaboration was measured as co-authorship involving different countries, while at the country and university levels, the count was based on the number of authors, regardless of whether partnering was national or international. We don't take in account double affiliation of authors.

Three attributes were calculated for the two output tiers: collaborative rate, co-authorship rate and average partner score. The first gives the percentage of co-authored papers (Gómez, Fernández, and Sebastian, 1999). The second measures the mean number of authors per paper. The third measures the degree of partnering, calculated as the number of papers with one author, two and so on. The partner score of an aggregate was found by assigning 0 points to papers with one author, 1 point to papers with two, 2 to papers with three, 3 to papers with four, 4 to papers with five and 5 to papers with more than five. The average partnering rate was obtained as the quotient between the partner score and the total number of co-authored papers (Levitt and Thelwall, 2009). The Spearman rank correlation coefficient (Spearman's rho) was calculated for the collaboration indicators in each output tier and two levels of analysis.

A relative indicator, the mean number of citations received per paper for each degree of collaboration (one, two, three, four, five or more than five authors), was used to determine the impact of partnered output for both output tiers (total and top 10%) and the two levels of analysis (countries and universities). Self-citations were included in the count because they have been shown to have no effect on the results when a large number of papers is analysed, as in the present case (Glänzel and Thijs, 2004, p. 286).

The Kruskal-Wallis test was used to ascertain whether the mean citation impact value varied with the size of the research team (variant factor) in the two output tiers and two levels of analysis (dependent variables). That test, the descriptive data and statistical analysis were conducted with SPSS 20 statistical software.

Results

World

Total output in the Scopus journal category "Library and Information Science" grew steadily over the period. In 2003-12, rose at a year-on-year rate of 9.8%. (Figure 1).

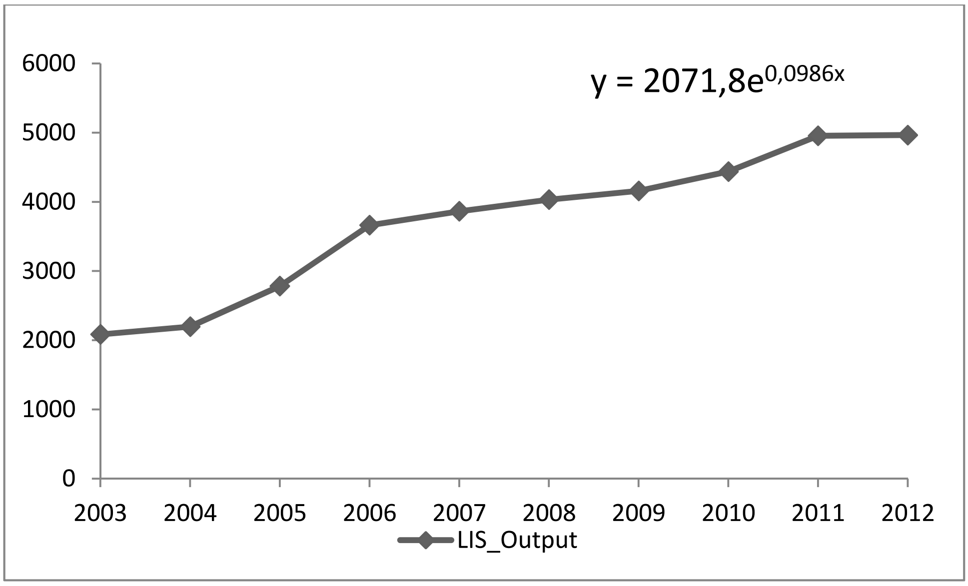


Figure 1. Scopus category Library and Information Science. Total output, 2003-12

The proportion of LIS articles with more than one country affiliation is analysed below based on two indicators: collaborative rate or proportion of internationally co-authored papers and co-authorship rate or mean number of countries involved per paper (Figure 2). The collaborative rate followed an upward trend (+13.4 %) throughout the period studied. In 2012, 62% of all papers originated from authors residing in different countries. Further to the co-authorship rate data, the median number of countries involved was 1.9: in other words, collaboration between researchers from two countries (dyadic relationships) prevailed. The mean number of countries per LIS manuscript grew across the period (+0.41).

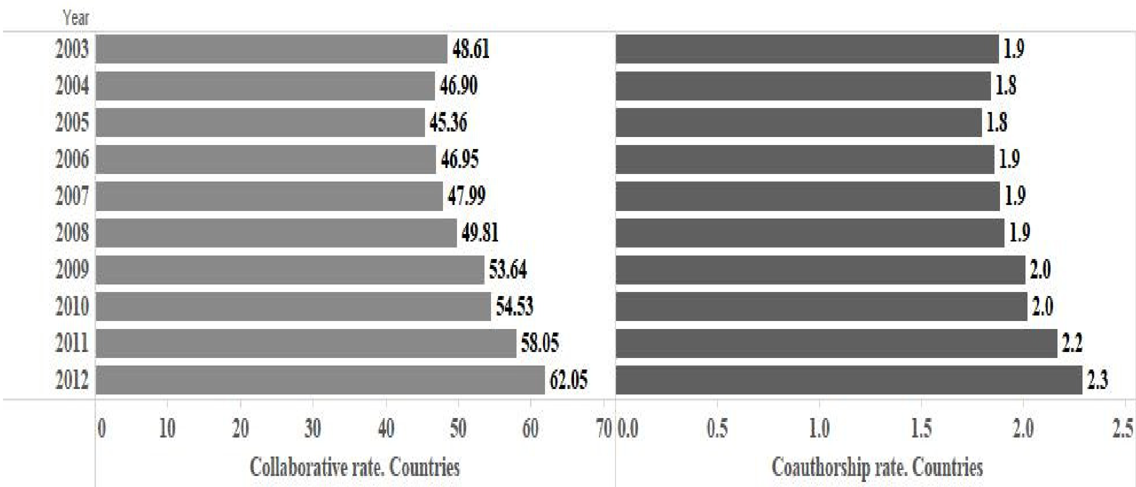


Figure 2. Collaborative and co-authorship rates. Scopus category Library and Information Science, 2003-2012

Figure 3 shows the pattern of international collaboration in LIS, measured as the number of country affiliations per paper (in per cent of total) across the period studied. The number of papers with a single country affiliation declined, while the number with two (+1.49 %) or three (+ 4.6 %) prevailed. The number of papers involving more than three countries has increased, although the growth rate varies: up 4.6 % for four countries, 1.94 % for five and 2.22 % for over five.

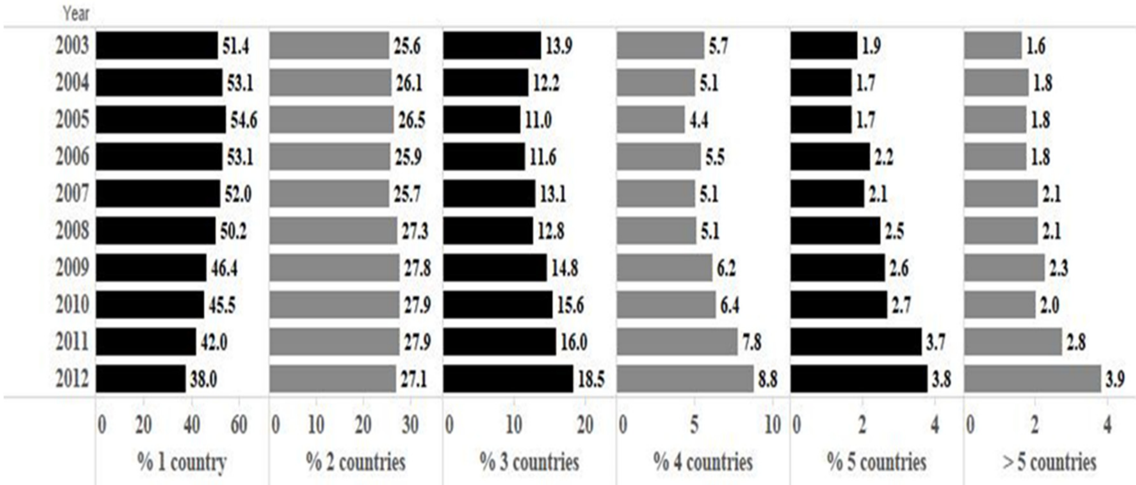


Figure 3. Collaborative levels of international collaboration. Scopus category Library and Information Science, 2003-2012

Taking the period as a whole, the values for all the degrees of collaboration in the "excellence" tier, expressed in per cent of the total (Table 1), were higher than for total output. To put it another way, authorship was more international in the more than in the less widely cited LIS papers. As the table shows, the differences in collaboration between the two types of output were substantial. More specifically, the collaborative rate was 42% higher and the co-authorship rate 32% higher in the top 10% tier than in the total.

Table 1. Degree of international collaboration, co-authorship and collaborative rates. Scopus category Library and Information Science. Total and top 10% output, 2003-12

Degree of collaboration	*2003-2012*	*Top 10% 2003-12*
% 1 country	47.5	25.11
% 2 countries	26.93	32.6
% 3 countries	14.34	20.04
% 4 countries	6.25	10.57
% 5 countries	2.65	6.17
% > 5 countries	2.32	5.51
Co-authorship rate	2	2.65
Collaborative rate	52.5	74.89

Countries

Table 2 gives the LIS category collaboration indicators for the 30 countries with an output of ≥150 papers between 2003 and 2012, ranked by the collaborative rate for the total (i.e., irrespective of the number of authors).The table gives the percentage of co-authored papers, the average partner score and the mean co-authorship rate.

Table 2. Collaborative rate, average partner score and co-authorship rate in LIS for the 30 most productive countries. 2003-2012 (according Total collaborative rate, ascending)

Country	*Total*	*Top 10 %*	*Collaborative rate*		*Average partner score*		*Co-authorship rate*	
			*Total**	*Top 10 %*	*All*	*Top 10 %*	*All*	*Top 10 %*
United States	16759	1772	45.56	72.8	1.84	1.91	1.89	2.48
Germany	1594	77	49.06	78.31	1.96	2.35	2.07	3.02
Austria	359	29	53.2	79.31	2.21	2.70	2.35	3.31
United Kingdom	3500	522	54.26	76.44	1.85	1.98	2.06	2.68
Denmark	276	237	57.97	66.67	1.74	1.70	2.07	2.14
Canada	1724	237	58.12	79.75	1.91	1.84	2.18	2.53
Israel	205	36	58.54	58.73	1.79	1.95	2.06	2.11
Norway	186	81	58.6	61.54	1.93	1.58	2.18	2.39
New Zealand	259	46	58.69	69.7	1.64	1.70	2.01	2.48
Australia	1189	155	59.71	77.42	1.85	1.98	2.13	2.5
India	598	46	60.87	73.91	1.62	1.85	2.05	2.98
Finland	357	77	61.06	64.94	1.76	1.58	2.2	2.08
Hungary	166	54	63.86	72.22	1.91	2.10	2.28	2.53
Sweden	288	70	63.89	72.22	1.99	2.12	2.42	3.04
Netherlands	797	90	65.12	76.83	1.89	1.87	2.31	2.61
Ireland	183	36	66.67	75	1.83	2.48	2.4	3.54
Belgium	528	144	67.05	78.47	2.00	2.04	2.4	2.52
France	758	66	67.94	86.36	2.07	2.14	2.51	3.34
Japan	473	46	68.5	76.09	2.38	2.37	2.72	2.91
Spain	1666	171	73.71	87.13	2.05	2.19	2.57	3.03
Italy	556	90	74.46	85.56	2.06	2.06	2.66	3.24
Brazil	562	28	75.8	100	1.87	2.79	2.44	3.66
South Korea	430	81	76.98	87.65	1.90	1.80	2.5	2.5
Iran	332	36	78.61	88.89	1.63	1.91	2.33	2.71
Taiwan	563	105	79.22	81.9	1.91	2.03	2.5	2.79
Switzerland	261	88	79.31	88.64	2.17	2.15	2.98	3.46
Greece	275	37	80.36	91.89	2.04	2.35	2.8	3.86
Singapore	288	70	80.56	81.43	1.93	2.00	2.57	2.68
Hong Kong	353	88	81.87	88.64	2.02	1.97	2.68	2.81
China	1483	237	90.02	92.41	2.50	2.40	3.34	3.48
Median	451.1	74.5	65.9	78.4	1.9	2.0	2.4	2.8
Mean	1232.3	161.7	67.0	79.0	1.9	2.1	2.4	2.8
SD	3015.70	319.93	10.96	9.60	0.20	0.30	0.32	0.47

Note: SD: Standard deviation; * Sort criterion

The countries with the greatest output were United States, United Kingdom, Canada, Spain, Germany, China and Australia, while the United States, United Kingdom, Canada, China, Denmark and Spain had the highest rates of widely cited papers.

Papers with a single author did not prevail in Asian countries such as China, Singapore, Korea, Taiwan or Japan. Nor was such authorship predominant in Mediterranean countries such as Greece, Italy, France or Spain, small European countries such as Belgium, Ireland or Switzerland or emerging countries such as Brazil.

In all countries as a whole, the median collaboration indicators were higher in the top 10% tier. In the US and UK, for instance, the collaborative rates for "excellence" papers were 60 and 41% higher than the values for the total, respectively, while the co-authorship rates were 31 and 30% higher.

Significant direct correlations were found among the collaboration indicators for total output (Table 3), while significance was particularly high for the top 10% (Table 4).

Table 3. Correlation among collaboration indicators

		Collaborative rate	Average partner score	Co-authorship rate
Collaborative rate	Spearman's rho	1	0.404*	0.860**
	Sig. (two tailed)	.	0.027	0.000
	N	30	30	30
Average partner score	Spearman's rho	0.404*	1	0.766**
	Sig. (two tailed)	0.027	.	0.000
	N	30	30	30
Co-authorship rate	Spearman's rho	0.860**	0.766**	1
	Sig. (two tailed)	0.000	0.000	.
	N	30	30	30

*. Correlation significant at 0.05 (two-tailed).

**. Correlation significant at 0.01 (two-tailed).

Table 4. Correlation among collaboration indicators for top 10% papers

		Collaborative rate, top 10%	Average partner score, top 10%	Co-authorship, top 10%
Collaborative rate top 10%	Spearman's rho	1	0.473**	0.660**
	Sig. (two tailed)	.	0.008	0.000
	N	30	30	30
Average partner score top 10%	Spearman's rho	0.473**	1	0.831**
	Sig. (two tailed)	0.008	.	0.000
	N	30	30	30
Co-authorship top 10%	Spearman's rho	0.660**	0.831**	1
	Sig. (two tailed)	0.000	0.000	.
	N	30	30	30

**. Correlation significant at 0.01 (two-tailed).

Universities

A total of 20 universities were identified that published at least 30 top 10% library and information science papers in 2003-12. Table 5 lists the total and "excellence" output and author collaboration indicators for each of these universities. The sort criterion chosen was collaborative rate, total.

Neither the number of papers nor the mean impact factor were normally distributed in each university. The median collaboration indicator values showed that the top 10% papers tended to involve more partnering and a larger mean number of co-authors. Nonetheless, the teams authoring the top 10% papers were not very large, for the average partner score was lower than for output as a whole. That finding was the result of the high percentage of top 10% papers written singly or by a small number of co-authors in universities with a large "excellence" output, such as Urbana-Champaign (46% of such papers were single-authored) Leiden (36%), Western Ontario (35%), Rutgers and Michigan, Ann Arbor (both with 28%). Hence the lack of a significant correlation between collaborative rate and average partner score in the top 10% output (Spearman's rho=0.2; p<0.01) and the negative correlation for total output (Spearman's rho = -0.2; p<0.01).

Table 5. Collaborative rate, average partner score and co-authorship rate in LIS for the 20 universities with the most influential output. 2003-2012 (according Total collaborative rate, ascending)

			Collaborative rate		*Average partner score*		*Co-authorship rate*	
University	*Total*	*Top 10%*	*Total**	*Top 10 %*	*Total*	*Top 10%*	*Total*	*Top 10%*
Rutgers, The State University of New Jersey	309	42	52.10	71.43	2.30	1.70	2.21	2.33
University of Illinois at Urbana-Champaign	428	37	56.31	54.05	2.03	1.60	2.13	1.97
University of Pittsburgh	244	55	61.48	85.45	2.28	2.47	2.51	3.6
Pennsylvania State University	340	71	61.76	85.92	1.78	1.79	2.17	2.69
University of Washington	271	43	63.84	76.74	2.46	2.42	2.8	3.32
University College London	167	34	64.67	82.35	2.31	3.00	2.49	3.68
University of Arizona	196	45	65.31	82.22	2.09	2.27	2.35	3
Indiana University-Bloomington	323	107	66.87	75.70	2.22	1.99	2.56	2.46
Leiden University	128	66	67.19	63.64	2.13	2.36	2.65	2.8
University of Amsterdam	175	95	67.43	80.00	1.56	1.50	2.13	2.31
University of Maryland, College Park	238	72	67.65	87.50	2.27	2.16	2.6	2.92
Florida State University, Tallahassee	211	40	68.72	75.00	2.24	1.87	2.66	2.39
University of Western Ontario	176	37	71.02	64.86	1.77	1.50	2.27	2.06
University of Michigan, Ann Arbor	205	50	71.71	72.00	2.43	2.36	2.87	2.9
Drexel University	198	43	72.22	81.40	2.09	2.06	2.63	3

University of North Carolina, Chapel Hill	361	64	72.30	78.13	2.57	2.22	3.06	2.95
University of Wolverhampton	132	68	80.30	83.82	1.57	1.54	2.25	2.21
Catholic University of Leuven	227	94	81.50	82.98	1.95	1.90	2.66	2.67
Nanyang Technological University	158	41	84.81	85.37	2.01	2.03	2.73	2.88
Universidad Granada	228	42	88.16	97.62	2.23	2.76	2.99	3.78
Median	219	47.5	67.5	80.7	2.2	2.0	2.58	2.84
Mean	235.75	57.30	69.27	78.31	2.11	2.07	2.54	2.80
SD	80.43	21.64	9.06	9.74	0.28	0.42	0.28	0.52

Note: SD: Standard deviation; *Sort criterion

Impact

Since the citations per paper by number of authors exhibited extreme variations and were biased at country and university levels and for both types of output, this parameter was compared among six independent groups (papers with 1, 2, 3, 4, 5 or >5 authors or countries) at both levels of analysis. The Kruskal-Wallis test was used to confirm the null hypothesis, i.e., that the research team size-based variations in the distribution of impact values, measured as citations per paper, were not significant.

At country level, the findings indicated the existence of significant inter-group (i.e., number of country affiliations) differences in the distribution of citations per paper for both total output ($p = 0.035 \leq \alpha = 0.05$) and the top 10% (0.00 α =0.05). At university level, the same Kruskal-Wallis test denoted significant inter-group (i.e., number of authors) differences in the distributions of citations per paper, both for the total ($p = 0.005 \leq \alpha=0.05$) and top 10% ($p = 0.033 \leq \alpha=0.05$) outputs. Consequently, at both levels of analysis, the mean impact of the papers published differed depending on the size of the research group involved.

Conclusions

This study, based on two levels of analysis and two output tiers, shows that international collaboration in library and information science, measured with the indicators described, grew throughout the period analysed, although more in terms of total output than of widely cited papers. Nonetheless, the data on international collaboration and the size of the research teams authoring widely cited papers reveal greater internationalisation in this tier than in output as a whole. Widely cited papers exhibit greater collaboration, higher co-authorship rates and greater international collaboration, for the worldwide total covers output from regions where partnering is more common than in North America, even in widely cited papers.

A more detailed analysis of country-level collaboration in which the unit analysed is all manner of personal collaboration shows that at this level papers are more collaborative and the number of authors per paper is larger. Partnering is much more intense in Asian countries (China, Hong-Kong, Singapore, Taiwan, South Korea, etc.), southern Europe (Greece, France, Spain, Italy), small European countries (Switzerland, Ireland) and emerging countries (Brazil, Iran) than in protestant or Anglo-Saxon European countries. More widely cited papers are more collaborative and in several countries the number of authors per “excellence” paper is larger than the number of authors per less widely cited paper.

In the period studied, the most widely cited LIS papers affiliated with the most influential universities were collaborative and had more than one author, although in some universities the most widely cited papers were singly authored. That notwithstanding, the conclusion drawn is that collaboration is more common in more than in less widely cited papers.

A non-parametric study of impact in the two levels of analysis, conducted to verify the existence of variations in citation frequency by degree of collaboration, yielded statistically significant differences.

Acknowledgments

The authors thank ScimagoLab for their support with compiling the data.

References

Adams, J., Gurney, K. A., and Marshall, S. (2007). Profiling citation impact: A new methodology. *Scientometrics*, 72, 325–344.

Bornmann. L., Moya-Anegón, F. de and Leydesdorff, L. (2012). The new excellence indicator in the World Report of the SCImago Institutions Rankings 2011. *Journal of Informetrics*, 6, 333-335.

Bridgstock, M. (1991). The quality of single and multiple authored papers; an unresolved problem. *Scientometrics*, 21, 37-48.

Craig Finlay, S., Ni, Ch. and Sugimoto, C.R. (2012). New methods for an old debate: utilizing reader response to investigate the relationship between collaboration and quality in academic journal articles. *Library and Information Science Research*, 34, 131-137.

Franceschet, M. and Costantini, A. (2010). The effect of scholar collaboration on impact and quality of academic papers. *Journal of Informetrics*, 4, 540-553.

Glänzel, W. (2001). National characteristics in international scientific co-authorship. *Scientometrics*, 51(1), 69–115.

Glänzel, W. (2002). Coauthorship patterns and trends in the sciences (1980-1998): a bibliometric study with implications for database indexing and search strategies. *Library Trends*, 50, 461-473.

Glänzel, W. and Thijs, B. (2004). The influence of author self-citations on bibliometric macro indicators. *Scientometrics*, 59, 281-310.

Gómez, I., Fernandez, M. T., and Sebastian, J. (1999). Analysis of the structure of international scientific cooperation networks through bibliometric indicators. *Scientometrics*, 44, 441–457.

Gorraiz, J., Reimann, R., and Gumpenberger, C. (2011), The importance of bilateral and multilateral differentiation in the assessment of international collaboration—A case study for Austria and six countries. In E. Noyons, P. Ngulube, and J. Leta (Eds.). *Proceedings of ISSI 2011—The 13th international conference on scientometrics and informetrics* (pp. 236–248). Durban, South Africa, 4–7 July 2011.

Hart, R.L. (2007). Collaboration and article quality in the literature of academic librarianship. *Journal of Academic Librarianship*, 33, 190-195.

Levitt, J.M. and Thelwall, M. (2009). Citation levels and collaboration within library and information science. *Journal of the American Society for Information Science and Technology*, 60, 434-442.

Persson, O., Glänzel, W., and Danell, R. (2004). Inflationary bibliometric values: The role of scientific collaboration and the need for relative indicators in evaluative studies. *Scientometrics*, 60, 421–432.

Rousseau, R. (2001). Are multi-authored articles cited more than single-authored ones? Are collaborations with authors from other countries more cited than collaborations within the country? A case study. In F Havemann, R. Wagner-Döbler, H. Kretschmer (Eds.). *Collaboration in Science* *and in Technology*. *Proceedings of the second Berlin workshop on Scientometrics and Informetrics* (pp. 173-176). Berlin: Gesellschaft für Wissenschaftsforschung.

Seglen, P. (1992). The Skewness of Science. *Journal of the American Society for Information Science*, 4, 628-638.

Waltman, L., Calero-Medina, C., Kosten, J. Noyons, E.C.M., Tijssen, R.J.W., van Eck, N.J. and Wouters, P. (2012) The Leiden Ranking 2011/2012: data collection, indicators and interpretation. *Journal of the American Society for Information Science and Technology*, 63, pp. 2419-2432.

Wuchty, S., Jones, B.F., and Uzzi, B. (2007). The increasing dominance of teams in production of knowledge. *Science*, 316, 1036-1038.

Influential Bloggers and Active Bloggers on ScienceNet: An Analysis of Popular Blogs

Chen Yue*, Zhang Liwei**, Wang Zhiqi, Liu Shengbo, Su Lixin and Hou Yu

WISE Lab, School of Public Administration and Law, Dalian University of Technology, Dalian 116085 China
**chenyuedlut@163.com, **taiyiliwei@126.com*

Abstract

This paper examines the influential bloggers and active bloggers on ScienceNet through the analysis of popular blogs to reflect the informal academic communication. Firstly, the top 20 most influential bloggers on ScienceNet are picked out by ranking the amount of being posted, being visited, being commented and being recommended respectively; Secondly, the top 102 most active bloggers are chosen by the social network analysis on the recommend bloggers in 1000 popular blogs. Analysis of blogger's informal academic communication will benefit to make an overall recognition to a researcher.

Keywords: ScienceNet; bloggers; SNA; informal academic communication

Introduction

Blog channel on ScienceNet has been increasingly popular for researchers to communicate and share academic ideas since 2002 in China. There has been more than 295 million bloggers until 2012, and will increase 73.1 million bloggers by a year in China[1]. The proportion of using blog in digital academic resources had up to 7% in 2008, according to digital scholarly communication[2]. Blog have been a informal academic communication way undoubtedly.

The ScienceNet is a leading website for academic circle in China, which is established by the Chinese Academy of Sciences, the Chinese Academy of Engineering, the National Natural Science Foundation, and hosted by the China Science Daily. The ScienceNet provides a wealth of academic information, which contains blog channel, conference channel, paper channel, the Meth Project channel, laboratory channel, etc. The blog channel is composed by several columns, including thesis communication, teach and research experiences, the popular science highlights, analysis of an issue, news blogs, etc. The ScienceNet carry out a real-name registration system, and the bloggers on ScienceNet are mostly composed by researchers, scientists, college teachers, technicists, and some other intellecturals. Bloggers should visit, comment, recommend, cooperate, communicate, discuss with each other in term of their academic interests and hobbies. Accordingly, some academic information will be spread and integrated to produce some new ideas

There are some deeper analysis on ScienceNet from the perspective of friend relations[3][4], the perspective of blog content[5][6][7] in recent years. However, there are few researches from the perspective of the relation network formed by visiting, commenting and recommending. In this paper, we select the popular blogs (Top highly recommended) on ScienceNet as data object to rank the influential bloggers and the active bloggers through an individual analysis and a network analysis respectively.

The influential bloggers by individual analysis

Blog Channel on ScienceNet is the main place for researchers in China to communicate or express opinions in an informal academic way. The influential bloggers should be powerful to disperse academic ideas and thoughts in an informal academic way. The powerful influence to bloggers should be embodied in the following three characteristics, including the blog content should be attractive to more people, propose more disputed topics, as well as be acknowledged.

We select three overall amount indicators, being visited, being commented and being recommended to reflect the three influential blogger's characteristics respectively.

4500 popular blogs are selected by the amount of being recommended from ScienceNet (Retrieval time is 2014-06-08). Different 20 influential bloggers listed in Tab. 1 are selected by ranking the three overall amount indicators. it shows that 'Rao Yi', 'Ji Shaocheng', 'Shi Yigong', 'Yu Hailiang' and 'Cao Guangfu' are most attractived with large amount of visitors, 'Rao Yi', 'Wu Feipeng', 'Li Xuekuan', 'Ji Shaocheng' and 'Cao Guangfu' prefer to propose more disputed topics with the lots of commences, meanwhile, 'Rao Yi', The ideas and thoughts in the blog of 'Wu Feipeng', 'Li Xuekuan', 'Ji Shaocheng', 'Meng Jin' and 'Huang Xiuqing' are acknowledged by more bloggers with a large numbers of recommends. The discordant results ranked by the three overall indicators indicate the various features of influential bloggers.

It is interesting that 'Rao Yi' and 'Shi Yigong' with a larger amount of being visited (3009242 and 1245794 respectively), being commented (13371 and 3684 respectively) and being recommended (15806 and 3998 respectively) compared with a fewer blogs being posted (113 and 17) respectively, which can be supported by the statistics data listed in column "per." , that is the rate of overall amount of being visited/ being commented/ being recommended to the overall amount of being posted by the blogger. Consequently 'Shi Yigong' and 'Rao Yi' produce the strongest influence among the bloggers on ScienceNet.

Tab. 1 The top 20 influential bloggers ranked by the three indicators

	Being posted		Being visited			Being commented			Being recommended		
No.	Blogger	amt.	Blogger	amt.	Per.	Blogger	amt.	Per.	Blogger	amt.	Per.
1	Li Xuekuan	213	Rao Yi	3009242	26630	Rao Yi	13371	118	Rao Yi	15806	140
2	Wu Feipeng	165	Ji Shaocheng	1259658	8075	Wu Feipeng	11298	68	Wu Feipeng	15140	92
3	Ji Shaocheng	156	Shi Yigong	1245794	73282	Li Xuekuan	10795	51	Li Xuekuan	14599	69
4	Meng Jin	136	Yu Hailiang	1226530	14098	Ji Shaocheng	10550	68	Ji Shaocheng	13517	87
5	Huang Xiuqing	121	Cao Guangfu	990049	8320	Cao Guangfu	9368	79	Meng Jin	10381	76
6	Cao Guangfu	119	Wu Feipeng	976145	5916	Yu Hailiang	7853	90	Huang Xiuqing	10260	85
7	Rao Yi	113	Xing Zhizhong	945393	10504	Huang Xiuqing	7769	64	Cao Guangfu	8737	73
8	Wu Yishan	108	Wen Shuangchun	875488	14839	Zeng Yongchun	6575	73	Wang Dehua	7726	80
9	Wang Dehua	96	Huang Xiuqing	846724	6998	Chen An	6504	73	Yu Hailiang	7699	88
10	Xing Zhizhong	90	Chen An	767350	8622	Meng Jin	5744	42	Wu Yishan	7299	68
11	Zeng Yongchun	90	Wang Feiyue	761096	69191	Wang Dehua	5371	56	Chen An	7146	80
12	Chen An	89	Wang Dehua	745175	7762	Lin Zhongxiang	5117	72	Jia Wei	7126	98
13	Yu Hailiang	87	Meng Jin	717568	5276	Xing Zhizhong	4939	55	Xing Zhizhong	7034	78
14	Jia Wei	73	Wang Hongfei	708024	16858	Jia Wei	4783	66	Zeng Yongchun	6811	76
15	Lin Zhongxiang	71	Li Xuekuan	677229	3179	Wen Shuangchun	4502	76	Lin Zhongxiang	6002	85
16	Wen Shuangchun	59	Jia Wei	675553	9254	Wu Baojun	4295	102	Wen Shuangchun	5476	93
17	Chen Xiangming	52	Cheng Daizhan	663829	26553	Wu Yishan	3902	36	Wu Baojun	4502	107
18	Peng Silong	51	Wang Yuncai	643353	14297	Shi Yigong	3684	217	Peng Silong	4271	84
19	Zhao Meidi	51	Wu Yishan	638964	5916	Wang Hongfei	3673	87	Cheng Daizhan	4267	171
20	Zhuang Shiyu	50	Lin Zhongxiang	632390	8907	Cheng Daizhan	3543	142	Shi Yigong	3998	235

Note: 'amt.'indicates the overall amount; 'Per'indicates the rate of overall amount of being visited/ being commented/ being recommended to the overall amount of being posted by the blogger.

Active bloggers by social network analysis

The active bloggers should appear more frequently on the ScienceNet. The active degree should not be only embodied in the overall amount of blogs posted, but also the frequency of recommend other blogs. In this paper, we pay more attention on the network influence to metric the active degree for bloggers.

Similar to the citation motivation in literature, recommendation means the acknowledgment to the ideas or thoughts showed in blog content. The co-occurrence network formed by recommend bloggers will present several blogger groups with different topic fields, core groups with active bloggers, as well as the relationship between different blogger groups on ScienceNet.

The top 1000 popular blogs are selected from 4500 popular blogs above mentioned, an 74958*74958 matrix are constructed with 74958 recommend blogger. We make a network (Fig. 2) with 699 notes and 49502 lines by omitting the bloggers that co-occurrence time is less than 30 for convenience and making sense.

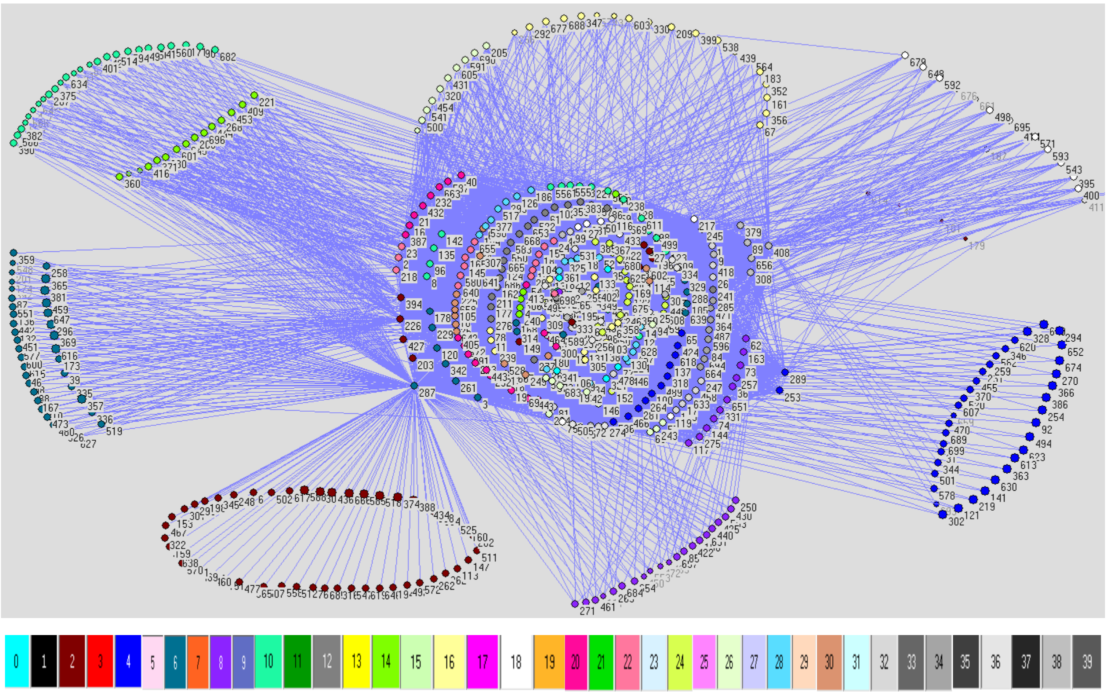


Fig. 2 The co-occurrencenetwork of 699 bloggers

39 partitions with different colors are showed in Fig. 2, and the 39 partitions represent 39 topics in mainly 4 interdisciplinary communication groups (Tab. 2) , It shows the strong interdisciplinary communication among Biology, Life Science, and Medical, among Physics, Engineering Science, Materials science, Chemistry, Organic Chemistry and Environmental chemistry, among Mathematics, Information Science, and Library and Information Science, as well as between Earth Science and Atmospheric Science. The forming of the four interdisciplinary communication groups is relevant to the disciplinary features.

Tab. 2 The main research areas of partitions of bloggers

No.	Partitions(topics)	Disciplinary fields
1	4、7、8、9、11、14、16、19、26、27、28、30	Biology; Life Science; Medical
2	0、3、5、10、17、21、22、23、24、25、32、36	Physics; Engineering Science; Materials science; Chemistry; Organic Chemistry; Environmental chemistry etc.
3	2、6、12、15、20、31、37、38、39	Mathematics; Information Science;Library and Information Science
4	1、13、18、29、33、34、35	Earth Science; Atmospheric Science

102 blogger nodes are selected as a blogger core from 699 blogger nodes presented in Fig. 2, and the blogger core group is just located in the center of Fig. 2. A co-occurrence network of 102 bloggers is showed in Fig. 3, and the co-occurrence frequency is more than 150 for each blogger nodes. These 102 bloggers should be seen as active bloggers, and they can be divided into 11 sub-groups with different colors in Fig. 3 by K-core[8][9][10] value. The blogger with blue nodes are most actively on ScienceNet.

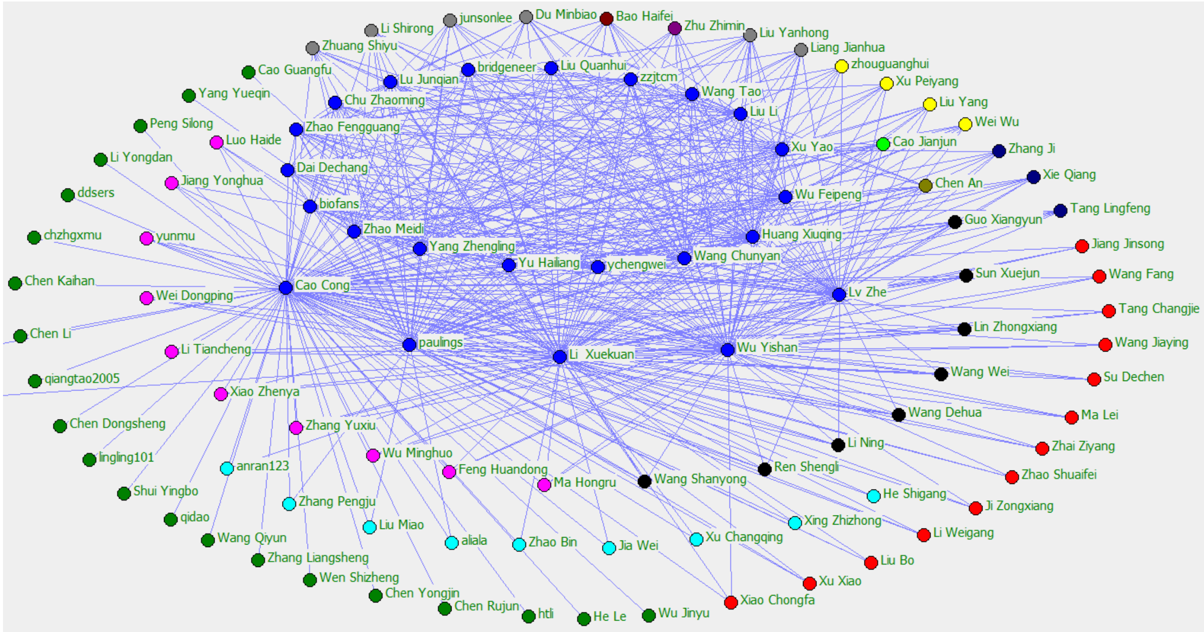


Fig. 3 The co-occurrence network of core blogger group

The role for a node in a network normally examined through centrality analysis in SNA. There are three indicators to measure the centrality for a node, including Degree Centrality, Betweenness Centrality, and Closeness Centrality. Degree Centrality is defined as the number of links incident upon a node. Betweenness Centrality quantifies the number of times a node acts as a bridge along the shortest path between two other nodes, It was introduced as a measure for quantifying the control of a human on the communication between other humans in a social network by Linton Freeman.[11] The farness of a node is defined as the sum of its distances to all other nodes, its closeness is defined as the inverse of the farness.[12][13] Tab.4 lists the top 30 active bloggers (blue blogger nodes in Fig. 3) according to the three Centrality measurement indicators. There are 28 same bloggers listed in the three centrality columns, and only little difference in ranking order. Consequently, the 30 bloggers are surely the most active on SciencNet no matter from local influence, connection role andwhole influence in the network.

Tab. 4 The active bloggers ranking by centrality measurement

No.	Ranking by degree		Ranking by betweenness		Ranking by farness	
	Blogger	Degree	Blogger	Betweenness	Blogger	Farness
1	Cao Cong	58117	Cao Cong	73343	Cao Cong	713
2	Li Xuekuan	45958	Li Xuekuan	21789	paulings	835
3	paulings	43805	paulings	21273	Li Xuekuan	837
4	Wu Yishan	42091	Wu Yishan	12511	Wu Yishan	875
5	Lv Zhe	40057	Lv Zhe	12271	Lv Zhe	882
6	Yang Zhengling	30963	Wang Tao	4644	Zhao Fengguang	966
7	Wang Chunyan	30256	biofans	4434	Wang Chunyan	977
8	Zhao Fengguang	29656	Wang Chunyan	4016	Yang Zhengling	980
9	Xu Yao	29304	Zhao Fengguang	4007	Xu Yao	980
10	Huang Xiuqing	29166	Xu Yao	3914	Wang Tao	985
11	Zhao Meidi	29039	Yu Hailiang	3830	Liu Li	988
12	Liu Li	28304	Yang Zhengling	3387	Yu Hailiang	991
13	Yu Hailiang	27283	Liu Li	3351	Zhao Meidi	1000
14	Wang Tao	27186	Du Minbiao	3102	Huang Xiuqing	1002
15	Lu Junxi	26675	Zhao Meidi	2980	Liu Quanhui	1026
16	Wu Feipeng	25726	Huang Xiuqing	2584	Du Minbiao	1031
17	Liu Quanhui	24495	Dai Dechang	2077	Lu Junxi	1034
18	Dai Dechang	24046	Lu Junxi	2015	Chu Zhaoming	1043
19	Du Minbiao	23715	Chu Zhaoming	1911	Wu Feipeng	1045
20	Chu Zhaoming	23246	Liu Quanhui	1887	Dai Dechang	1050
21	zzjtcm	23212	Liang Jianhua	1763	Liang Jianhua	1052
22	biofans	22716	Wu Feipeng	1707	bridgeneer	1065
23	bridgeneer	22034	ychengwei	1608	biofans	1066
24	ychengwei	22004	Cao Jianjun	1568	ychengwei	1072
25	Liang Jianhua	21847	Xu Peiyang	1451	zzjtcm	1074
26	Li Turong	19847	bridgeneer	1099	Xu Peiyang	1077
27	Zhu Zhimin	19806	zhouguanghui	1072	Zhu Zhimin	1085
28	Chen An	19692	zzjtcm	938	Cao Jianjun	1086
29	Cao Jianjun	19417	Zhu Zhimin	826	zhouguanghui	1091
30	Bao Haifei	19297	Li Turong	796	Li Turong	1096

Conclusions and discussions

The way of research is increasingly changed by the rapid development of internet social media, and the informal academic communication, such as blog or twitter, etc., is increasingly become the supplement to formal academic communication, such as publication or teaching, etc. This paper **examines** the influential bloggers through individual analysis and active bloggers through network analysis on the basis of the popular blogs on ScienceNet.

References

[1] CNNIC. (2012). Handbook on Internet Protocol (IP)-Based Networks and Related. http://www.cnnic.cn/hiwfzyj/hwxzbg/hiwtjbg/201301/t20130115_38508.htm

[2] Nancy L, Maron k, Kirby Smith. (2011). Current models of digital scholarly communication: Results of an investigation conducted by Ithaka Strategic Services for the Association of Research Libraries. Washington, DC: Association of Research Libraries. http://www.arl.org/bm~doc/current-models-report.pdf.

[3] Zhou, C., & Zhu, X. (2013). Study on LIS Blogs in Science Net.Documentation. *Information& Knowledge*, (5), 98-105.

[4] Zhou, C.(2012). Study on Sciencenet Blog Evaluation with Link Content Analysis. *The Library and Information Knowledge*, (4), 11-17.

[5] Qiu, J.,& Yu F. (2011).Research on the Author Distribution of Academic Information on the Internet.*Library and Information Service*,55(20), 15-19.

[6] Qiu, J.,&Wang, F. (2011).Analysis about knowledge Communication Network Based on Friends Link in the Blog Community--A Case Study of the Blog Community in www.sciencenet.cn.*The Library and Information Knowledge*, (6), 25-33.

[7] Li, C., Li, C., Han, M., &Duan, W. (2012).Analysis on the Current Situation of Academic Information Resources in the Blog Posts on ScienceNet. Library Tribune,3(22), 5-8.

[8] Baxter G J, Dorogovtsev S N, Goltsev A V, et al, (2012). K-Core Organization in Complex Networks [M]//Handbook of Optimization in Complex Networks. Springer US, 2012: 229-252.

[9] Bollabás B (1984). The evolution of sparse graphs[C]//Graph Theory and Combinatorics: Proc. Cambridge Combinatorial Conf. in honor of Paul Erdös (B. Bollabás, Ed.). 1984: 35-57.

[10] Seidman S B (1983). Network structure and minimum degree[J]. Social networks, 1983, 5(3): 269-287.

[11] Freeman, Linton (1977). "A set of measures of centrality based upon betweenness". Sociometry 40: 35–41. doi:10.2307/3033543.

[12] Alex Bavelas. Communication patterns in task-oriented groups. J. Acoust. Soc. Am, 22(6):725–730, 1950.

[13] Sabidussi, G. (1966) The centrality index of a graph. Psychometrika 31, 581–603.

An Analysis of Collaboration Pattern of Indian S&T Papers (Published during 2005-09)

Divya Srivastava, Arvind Singh Kushwah and Mona Gupta

Division of Publication & Information, Indian Council of Medical Research, New Delhi, India
drdivya.srivastava@gmail.com

Studying research collaborations has evolved into a major focus of bibliometrics and receives increasing attention from policy-makers and more general users. Modern research is regarded as increasingly complex and specialized, making it impossible for an individual researcher to master all the knowledge and technical skills needed. In a collaboration, different skills complement each other and this complementarity is hoped to stimulate knowledge sharing and the generation of innovation and new ideas. As a result, collaborative research activities do not only enable the pooling and sharing of resources for enhanced efficiency but are also linked to the quality of the research outcome.

This growing interest in research collaborations is also reflected in research funding programs. Grants awarded by many different funding institutions and for many different disciplines often seek to encourage – and at times require as a condition – collaborations between different countries, research fields or institutions. Being able to map and analyze research networks and collaboration has therefore evolved into a key issue for the design and assessment of research policies and related funding programs. Collaboration is now actively promoted with a view to breaking down the barriers between research institutions, industry, commerce, government and the public services. Specific driving factors include: the growth of the knowledge economy and attempts to strengthen the economic and social contribution of research Collaboration occurs at various levels including individuals, groups, departments, institutions, sectors and countries. The latter may emerge from political memoranda of understanding between nations, although definitions of higher levels of collaboration are no easier to arrive at than for inter-individual collaboration. Nevertheless, it is important to make this distinction between the different levels because an inter-institutional or international collaboration may not necessarily entail an inter-individual collaboration. What constitutes a collaboration varies across institutions, fields, sectors and countries, and changes with time. Some collaboration is formal, much more is informal.

Smith was one of the first researchers to observe an increase in the incidence of multiple-author papers and to suggest that such papers could be used as a proxy measure for collaboration among groups of researchers. In the present study explicit networks of such connections have been constructed by using data drawn from SCOPUS, for Indian Papers during the period of 2005-09. The results include distribution of numbers of collaborators of authors, demonstrate the presence of clustering in the networks, and highlight a number of apparent differences in the patterns of collaboration between the fields studied. Differences among 'institutions' have been investigated with regard to productivity, number and rate of national/international collaborations. Here we refer to collaboration in terms of intra- national & international collaboration. The degree of (national, international) collaboration is the percentage of collaborative articles out of the total number of articles. All Indian institutional addresses on the papers were unified to a set of standard institutional names and each standard name was assigned to an institutional sector. Using co-publication as unit of analysis bibliometrical studies have for example empirically demonstrated that researchers collaborate more than ever.

Data sources for Collaboration Analysis – advantages and disadvantages

Any collaboration analysis is primarily based on the SCI produced by the Institute for Scientific Information (ISI) in Philadelphia, USA or SCOPUS produced by Elsevier. The first advantage is that both the services cover all science fields. This is a necessity if one is looking at whole research systems. In addition, their coverage is unambiguous because every item from every journal is indexed. Coverage in other databases is ambiguous for indicator purposes because although they include all items from core journals, only items considered relevant to the subject of the database are included from secondary journals.

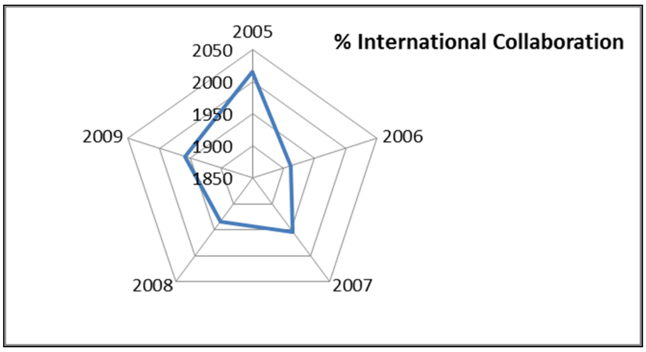
The second advantage is that all author addresses listed on the paper are included in these Databases. This is a necessity for studying institutional output as collaboration is so extensive. Only first addresses are included in other databases, and so papers on which an institution's address was not listed first cannot be credited to the institution. This source of error is substantial and growing as the rate of institutional collaboration increases. Only the first address is needed to contact authors of a paper, so listing only the first address is not a problem from the perspective of scientists searching the literature. From the policy perspective, the address that happens to be listed first is a social artifact and not of great policy interest in comparison to the total output of the institution. Of course, only if all addresses are listed can collaboration be studied.

The third advantage is that total number of references is also included in these databases. Citation counts can be derived from these references and used as a partial indicator of the impact previous research has had on succeeding work Citation counts are such a useful adjunct to policy analysis that almost by themselves their presence justifies using this data for policy analysis.

In our research we have used bibliometrics principals to study international collaboration patterns from an evolutionary point of view. Using co-authorship we have explored the collaboration of researchers in India with those in other countries (international) and with Indian Institutions (intra-national) We have examined how collaboration patterns are different among the different broad subject areas. In the present study, for India the share of publications with multiple institutions represented on them grew from 40% to 61% between 2005 and 2009.

Collaboration- International (Source: SCOPUS 2005-09)

The pattern of international collaboration has indicated that maximum international collaboration occurred during the year of 2005 followed by 2009, 2007, 2008 and finally 2006 in the same order.



At international level the data has indicated that a total of 159 countries have collaborated with an Indian authored paper. This may be noted that more than one country may have appeared in the same paper. The collaborating countries are distributed around the globe (2005-09).

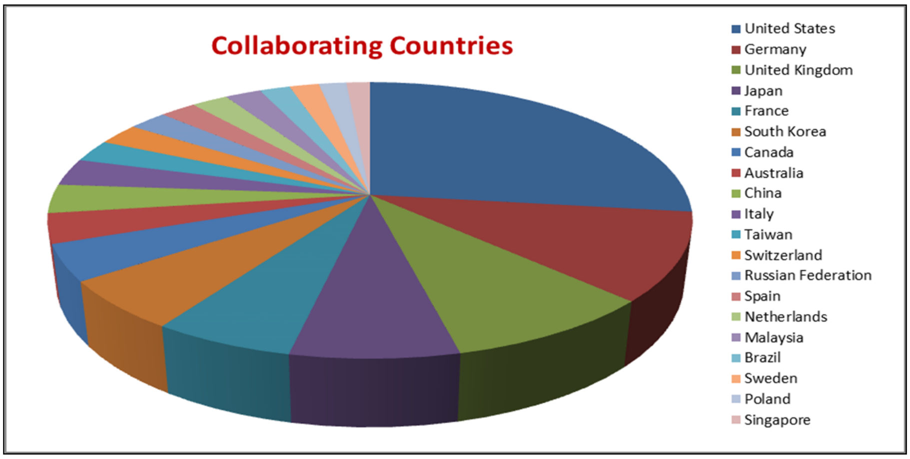
S No.	Collaborat. Country	Total papers in Collabor at.
1	Afghanistan	11
2	Albania	6
3	Algeria	27
4	Argentina	318
5	Armenia	29
6	Australia	1,521
7	Austria	461
8	Azerbaijan	10
9	Bahrain	22
10	Bangladesh	200
11	Barbados	5
12	Belarus	23
13	Belgium	497
14	Benin	9
15	Bhutan	17
16	Bolivia	11
17	Bosnia and Herzegovina	3
18	Botswana	21
19	Brazil	798
20	British Indian Ocean Territory	3
21	Brunei Darussalam	3
22	Bulgaria	129
23	Burkina Faso	4
24	Burundi	2
25	Cambodia	12
26	Cameroon	22
27	Canada	1,857
28	Cayman Islands	6
29	Chile	148
30	China	1,477
31	Colombia	228
32	Congo	8
33	Costa Rica	24
34	Cote d'Ivoire	4
35	Croatia	120
36	Cuba	20
37	Cyprus	37
38	Czech Republic	501
39	Democratic Republic Congo	1
40	Denmark	328
41	Dominican Republic	5
42	Ecuador	183
43	Egypt	145
44	El Salvador	1
45	Eritrea	24
46	Estonia	42
47	Ethiopia	140
48	Fiji	59
49	Finland	347
50	France	2,625
51	Gabon	5
52	Georgia	21
53	Germany	4,504
54	Ghana	25
55	Greece	218
56	Guadeloupe	5
57	Guatemala	6
58	Guinea	2
59	Guyana	2
60	Hong Kong	288
61	Hungary	312
62	Iceland	12
63	Indonesia	104
64	Iran	363
65	Iraq	26
66	Ireland	358
67	Israel	434
68	Italy	1,390
69	Jamaica	8
70	Japan	3,118
71	Jordan	51
72	Kazakhstan	24
73	Kenya	83
74	Kuwait	131
75	Kyrgyzstan	8
76	Laos	8
77	Latvia	12
78	Lebanon	23
79	Lesotho	4
80	Libyan Arab Jamahiriya	43
81	Liechtenstein	1
82	Lithuania	14
83	Luxembourg	5
84	Macedonia	4
85	Madagascar	11
86	Malawi	10
87	Malaysia	950
88	Maldives	7
89	Mali	13
90	Malta	7
91	Mauritius	11
92	Mexico	495
93	Mongolia	13
94	Montenegro	6
95	Morocco	24
96	Mozambique	9
97	Myanmar	28
98	Namibia	8
99	Nepal	252
100	Netherlands	952
101	New Caledonia	3
102	New Zealand	141
103	Niger	12
104	Nigeria	109
105	North Korea	2
106	Norway	193
107	Oman	119
108	Pakistan	161
109	Palestine	6
110	Panama	12
111	Papua New Guinea	11
112	Paraguay	2
113	Peru	33
114	Philippines	164
115	Poland	723

116	Portugal	370
117	Puerto Rico	56
118	Qatar	9
119	Romania	183
120	Russian Federation	1,011
121	Rwanda	4
122	Saint Lucia	3

123	Saudi Arabia	242
124	Senegal	6
125	Serbia	45
126	Sierra Leone	2
127	Singapore	622
128	Slovakia	127
129	Slovenia	236
130	South Africa	406

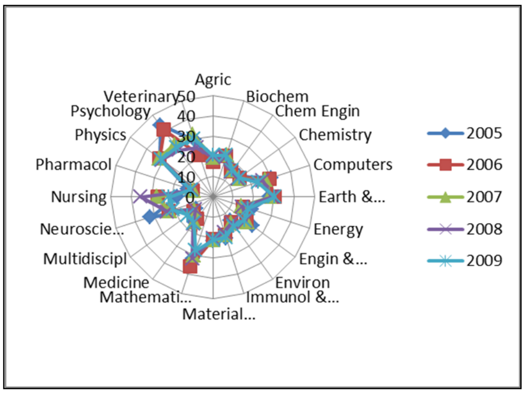
131	South Korea	2,538
132	Spain	956
133	Sri Lanka	133
134	Sudan	21
135	Sweden	776
136	Switzerland	1,105

The top most 20 countries as a co-authors were United States, Germany, United Kingdom, Japan, France, South Korea, Canada, Australia, China, Italy, Taiwan, Switzerland, Russian Federation, Spain, Netherlands, Malaysia, Brazil, Sweden, Poland and Singapore in the same order.



Collaboration- International (Subject Area wise)

We have also found significant differences in collaboration patterns across fields, through the analysis of 'Collaboration Index' (CI), Mathematics, Agricultural Sciences, Engineering, Computing & Technology is the field with the highest level of national publications whereas Biological Sciences, Medical Sciences, Physical Sciences, Chemical sciences and Earth & Atmospheric Sciences including Environmental studies are the fields with the highest level of both intra-national and international collaborations. These are the areas of scientific disciplines where teamwork and collaboration was priority during 2005-09. This characteristic can be considered as an indicator of maturity of research teams. In India, Medicine is a scientific discipline with a great tradition that counts with numerous well-established, internationally recognized teams International collaboration of these groups has been favored by the great diversity and complexity of Indian territory that has arouse the interest of many foreign scientists, mainly Americans, French, German and British. International Collaboration Index is a relative measure of the collaborative activity of a country.



Trend of International Collaboration in different Subject areas (2005-09)

Any bibliometric analysis for the comparisons between fields also need to be done with caution. Differences here may be due to different publication strategies Mathematicians and other theoreticians tend to publish fewer papers than researchers in experimentally-intensive fields such as life sciences, biological sciences or medical sciences Publications in the field of medicine do on average involve more authors than papers in the field of chemistry, agricultural or engineering. Technological scientists have on average a lower publication rate than others since the way of communicating research results often involves other means such as working papers, presentations or books etc. The top 15 Subject areas were covered by more than 50 % of papers. This indicates a clear shift towards more applied and frontline areas.

Analysis of Institutions having Collaboration

The bibliometric evidence indicates that the majority of scientific research is collaborative. The size and geographical location of an institution influences its collaboration profile. An institution and its researchers do not work in isolation; they work within broad and extensive research networks. The implications of the findings suggest that evaluation activities may need to make adjustments for the non-linear effect of institutional size when making comparisons. They also suggest we need a better understanding of how much collaboration policy actually influences a science system where the emerging nature and culture of scientific research appears to encourage collaborative activity.

We carried out an analysis of National Collaboration (% NC) & International Collaboration (% IC) also, institution's output ratio produced in collaboration with national & foreign institutions. The values are computed by analyzing an institution's output whose affiliations include more than one institution or country address. All Indian institutional addresses on the papers were unified to a set of standard institutional names and each standard name was assigned to an institutional sector. The bibliometric analysis involved examining the patterns of various types of collaboration in different S&T fields. There are a total of 5801 institutions contributing the total of 343895 papers. Out of this, ≥ 91.9% papers were produced as co-authored papers. A total of 50% papers were having co-authors from other institutions, national as well as international. The top most 150 institutes for all the five years in terms of 'international collaboration' are:

Rank	Institute	% IC
1	Mangalore University	63.47
2	Inter-University Centre for Astronomy and Astrophysics	57 82
3	Tata Institute of Fundamental Research	53.34
4	IBM India Research Laboratory	49.27
5	Institute of Mathematical Sciences	44.76
6	Institute of Physics Bhubaneswar	41.65
7	Variable Energy Cyclotron Centre	41.42
8	Raman Research Institute	39.74
9	Physical Research Laboratory	38.44
10	Inter-University Accelerator Centre	37.71
11	Harish Chandra Research Institute	37.5
12	Bharathiar University	36.48
13	Panjab University	34.43
14	Bharathidasan University	30.48
15	UGC-DAE Consortium for Scientific Research Indore	30
16	Saha Institute of Nuclear Physics	29.41
17	Indian Association for the Cultivation of Science	28.52
18	University of Hyderabad	27.99
19	University of Mysore	27.76
20	North-Eastern Hill University	27.6
21	Indian Institute of Technology, Bombay	27.54
22	S.N. Bose National Centre for Basic Sciences	27.07
23	Indian Institute of Technology, Kanpur	27.01
24	Indian Statistical Institute	26.13
25	Jawaharlal Nehru Centre for Advanced Scientific Research	26.01
26	National Geophysical Research Institute (sub)	25.91
27	Centre for Cellular and Molecular Biology (sub)	25.89
28	Sri Venkateswara University	25.64
29	University of Jammu	25.35
30	Indian Institute of Science	25.22
31	National Institute of Oceanography (sub)	24.66
32	University of Calcutta	23.56
33	University of Pune	23.53
34	Jamia Millia Islamia Central University	23.43
35	Madurai Kamaraj University	23.15
36	Indian Institute of Chemical Biology (sub)	22.73
37	Raja Ramanna Centre for Advanced Technology	22.65
38	Jawaharlal Nehru University	22.34
39	Bose Institute Kolkata	22.03
40	University of Delhi	21.87
41	Bengal Engineering and Science University, Shibpur	21.77
42	Jadavpur University	21.6
43	VIT University	21.53
44	Indian Institute of Technology, Delhi	21.45
45	Guru Nanak Dev University	21.32
46	National Institute for Interdisciplinary Science and Technology (sub)	21.14
47	Pondicherry University	21.04
48	Indian Institute of Technology, Madras	20.86
49	International Institute of Information Technology, Hyderabad	20.29
50	Devi Ahilya University	20.28
51	University of Burdwan	19.94
52	Indian Council of Medical Research	19.87
53	National Chemical Laboratory (sub)	19.64
54	Christian Medical College, Vellore	19.57
55	Birla Institute of Technology	19.38
56	National Physical Laboratory India (sub)	19.37
57	Tata Sons Ltd.	19.25
58	University of Madras	19.1
59	Bhabha Atomic Research Centre	18.48

60	Indian Institute of Technology, Kharagpur	18.45
61	National Institute of Mental Health and Neuro Sciences	18.45
62	Aligarh Muslim University	18.39
63	Indian Institute of Technology, Roorkee	18.22
64	National Institute of Technology Karnataka	18.18
65	Institute of Genomics and Integrative Biology (sub)	18.04
66	Tata Memorial Centre	17.92
67	Central Electrochemical Research Institute (sub)	17.79
68	Visva-Bharati University	17.74
69	Birla Institute of Technology and Science	17.67
70	Indian Institute of Technology, Guwahati	17.53
71	Banaras Hindu University	17.26
72	Cochin University of Science and Technology	17.17
73	Central Glass and Ceramic Research Institute (sub)	17.02
74	Karnatak University	16.91
75	Anna University	16.82
76	National Metallurgical Laboratory (sub)	16.73
77	Tamil Nadu Agricultural University	16.67
78	National Institute of Pharmaceutical Education and Research	16.38
79	University of North Bengal	16.33
80	Andhra University	16.01
81	Jamia Hamdard University	15.66
82	National Institute of Technology, Tiruchirappalli	15.39
83	Council of Scientific and Industrial Research*	15.29
84	University of Kalyani	15.24
85	Indira Gandhi Centre for Atomic Research	15.15
86	Sri Ramaswamy Memorial University	14.74
87	Shivaji University	14.54

88	Indian Space Research Organization	14.54
89	Chhatrapati Shahuji Maharaj Medical University	14.39
90	University of Rajasthan	14.38
91	Institute of Minerals and Materials Technology (sub)	14.21
92	National Institute of Technology Durgapur	13.24
93	Central Leather Research Institute (sub)	13.22
94	National Institute of Technology Rourkela	12.89
95	All India Institute of Medical Sciences	12.86
96	Motilal Nehru National Institute Of Technology	12.75
97	University of Lucknow	12.64
98	Gulbarga University	12.62
99	Manipal University	12.53
100	National Environmental Engineering Research Institute (sub)	12.45
101	Jawaharlal Nehru Technological University, Hyderabad	12.28
102	University of Mumbai	12
103	Jai Narain Vyas University	11.86
104	Chaudhary Charan Singh Haryana Agricultural University	11 72
105	Tezpur University	11.59
106	Punjab Agricultural University	11.53
107	Indian Institute of Toxicology Research (sub)	11.48
108	Bangalore University	11.45
109	University of Kerala	11.44
110	Seth Gordhandas Sunderdas Medical College and	11 05
111	Annamalai University	11.02
112	National Botanical Research Institute (sub)	11.02
113	Sri Siva Subramania Nadar College of Engineering	10.77

114	Govind Ballabh Pant University of Agriculture and Technology	10.7
115	Indian Institute of Chemical Technology (sub)	10.51
116	Sanjay Gandhi Postgraduate Institute of Medical Sciences	10.32
117	Osmania University	10 12
118	Sree Chitra Tirunal Institute for Medical Sciences and Technology	10 12
119	Shanmugha Arts, Science, Technology and Research Academy	
120	Indian Council of Agricultural Research	9.92
121	Rashtrasant Tukadoji Maharaj Nagpur University	9.92
122	Indian School of Mines	9.9
123	Postgraduate Institute of Medical Education and Research	9.77
124	Thiagarajar College of Engineering	9.75
125	National Institute of Technology Warangal	9.65
126	Allahabad University	9.46
127	Sardar Patel University	9.27
128	Kurukshetra University	9.09
129	Central Drug Research Institute (sub)	9.08
130	Central Food Technological Research Institute (sub)	8.82
131	Central Salt and Marine Chemicals Research Institute (sub)	8.47
132	The Maharaja Sayajirao University of Baroda	8.46
133	Kakatiya University	8.32
134	Kuvempu University	8.02
135	Thapar University	7.96
136	Himachal Pradesh University	7.91
137	Mohan Lal Sukhadia University	7.62
138	Vidyasagar University	7.57
139	Defence Research and Development Organisation	7.51
140	Indian Veterinary Research Institute	7.51
141	Guru Jambheshwar University of Science and Technology	7.32
142	Doctor Harisingh Gour University	7.14
143	Government Medical College and Hospital	7.13
144	Punjabi University	7.06
145	Dr. Babasaheb Ambedkar Marathwada University	6.64
146	University College of Medical Sciences	6.42
147	PSG College of Technology	6.37
148	Jawaharlal Institute of Postgraduate Medical Education and Research	5.83
149	Guru Tegh Bahadur Hospital	5.83
150	Tamil Nadu Veterinary and Animal Sciences University	5.29

This use of bibliometric has yielded some important insights: We find that researchers in smaller institutions co-author more with other intra-institutions than bigger institutes, while the international co-authorship rate is not dependent on the ‘status’ of the research institutes. It is interesting to note that some of the premier institutions from the sciences & technology field of India *eg*. Indian Institute of Science (30), Bhabha Atomic Research Centre (59), All India Institute of Medical Sciences(95), Indian Institute of Chemical Technology (115), Sanjay Gandhi Postgraduate Institute of Medical Sciences (116), Sree Chitra Tirunal Institute for Medical Sciences and Technology (118), Postgraduate Institute of Medical Education and Research (123), Central Drug Research Institute (129), Central Food Technological Research Institute (130), and Central Salt and Marine Chemicals Research Institute (131) are quite far off from the ranking list of the institutions in terms of ‘international collaboration. We also have analysed the pattern of 'intra-national collaboration' of institutes. Here the ranking have changed very significantly:

The top most 25 institutes in terms of 'intra-national collaboration were: Centre for Cellular and Molecular Biology, Variable Energy Cyclotron Centre, Indian Institute of Toxicology Research, Jawaharlal Nehru Centre for Advanced Scientific Research, Inter-University Centre for Astronomy and Astrophysics, Central Salt and Marine Chemicals Research Institute, Institute of Physics Bhubaneswar, National Institute for Interdisciplinary Science and Technology, Harish Chandra Research Institute, Panjab University, National Institute of Pharmaceutical Education and Research, University of Jammu, Indian Institute of Technology, Bombay, Birla Institute of Technology and Science, IBM India Research Laboratory, Indian Institute of Technology, Guwahati, Indian Institute of Technology, Delhi, Indian Institute of Technology, Kanpur, National Chemical Laboratory, Shivaji University, Tata Institute of Fundamental Research, Indian Institute of Technology, Kharagpur, Indian Institute of Technology, Roorkee, Indian Institute of Science and Indian Association for the Cultivation of Science.

Evidence from the study provides the following headline findings

- Collaboration is an essential feature of the research base. Collaboration in research is pervasive throughout. In all the papers the basic building block is inter-personal collaboration. They are based on individual researchers, who work collaboratively in a climate of shared intellectual interest and trust.
- Although, the inter-institutional or international collaboration may not necessarily entail inter-personal collaboration, evidence from the case studies suggests that a strong collaborative research base is an important success factor in the operationalization of higher levels of aggregation
- A greater proportion of publications from smaller institutions than from larger institutions involved domestic, intra-city, inter-city collaboration. On the other hand a greater proportion of the papers from larger institutions were having international collaboration than from smaller institutions.
- In India, the study revealed that Southern States were working more in collaboration with each other as compared to Northern States but later years are indicating a shift towards Northern Region.

Finally, with regard to mapping collaborations it is important to keep in mind that much collaboration do not result in co-published papers but may involve the sharing of research infrastructure, exchange of material or samples or some kind of informal collaboration which involve knowledge stimulation Using this policy makers will be able to understand the underlying factors and the cognitive behavior of researchers qualitative methods are needed to complement bibliometric analysis.

References

Bibliometric Study of India's Research Output and International Collaboration. A report, undertaken by Evidence, a Thomson Reuters business, on behalf of Research Councils UK (2011) Source: http://www.nature.com/news/2009/090708/full/news.2009.652.html and http://www.dst.gov.in/about_us/ar07?08/ar07?08index.htm.

Coordination and Collaboration in European Research towards Healthy and Safe Nanomaterials. (2011) J. Phys.: Conf. Ser. 304; 012001, (http://iopscience.iop.org/1742-6596/304/1/012001).

Elsevier Integrates New Collaboration Analysis Capabilities into Sci Val Strata to Further Enhance Research Evaluation: Research Trends (SCOPUS) September, 2011.

Evidence from the latest BIS report indicates that international collaboration on articles boosts impact through citations and adds to the UK's position as a 'world-class' research nation. blogs.lse.ac.uk/impactofsocialsciences/2011/11/14/evidence-from-the-latest-bis-report/ Accessed on 20th Nov. 2012.

Gupta B.M., Dhawan S.M. & Osareh, F. (2004). India-Middle-East Collaboration in S&T: an analysis through co-authored publication, *Library Herald*, 42 (4): 309-323.

Haustein S., Tunger D., Heinrichs G. & Baelz G. (2010). Reasons for and developments in international scientific collaboration: does an Asia-Pacific research area exist from a bibliometric point of view? *Scientometrics*, 86(3), 727-746.

Impact of Indian S&T Research Papers – Published during 2005-09: through Citation Analysis

Divya Srivastava, Arvind Singh Kushwah and Mona Gupta

Division of Publication & Information, Indian Council of Medical Research, New Delhi, India
drdivya.srivastava@gmail.com

Citations are used to measure impact. The premise underlying this indicator is that a research finding frequently referenced by other researchers has had greater impact on the research community than an infrequently cited paper. Impact is not the same as quality. However, in many instances impact and quality may be congruent. On the other hand, a contentious research finding, for example in Indian context, the claim of 'Therapy through Stem Cell', may be highly cited not because the work was of high quality but because it stimulated a vibrant debate about a research claim. In other words, it impacted the research community. We must never forget that negative impact can spawn new research ideas.

The simplest measure of impact is citations per paper. The choice of the citation window width is somewhat arbitrary. Typically, within five years most papers will receive about 40-50% of their citations. Narin has shown that the citation peak usually occurs in the second or third year after publication although this can vary across science fields. The narrower citation window provides a measure of the impact of faster moving, perhaps leading edge and research. However, one must keep in mind that the citation culture can vary from field to field and in some areas of research the rate of diffusion of new research findings can be much slower than in others *e.g.* Mathematics or Engineering.

Another factor to consider is the effect of self-citation (i.e. an author citing previously published work in a current paper) on the impact measure. Removing the effect of self-citation in a large corpus of publications is computationally difficult so the effect of self-citations is rarely considered. However, it has been demonstrated that for a large cohort of papers, such as those from a institution or a broad subject area, the percentage of self-citations remains fairly constant thus affecting the 'impact' in a similar and comparable manner across most institutions and subject area.

Table 1. Total Citations to the Papers Published during 2005-09

Year	*Total Papers*	*Total Citations*	*Self-Citation*	*Citation /paper*	*Cited Papers*
2005	28187	320648	109.6	15.84	20242
2006	31652	328389	113.859	10.23	32093
2007	36919	310326	107.357	10.15	30586
2008	43373	266414	92.238	7.34	36309
2009	44775	214443	76.567	5.72	37485

Source Data: WoS Expanded Online. Searched on March 2013

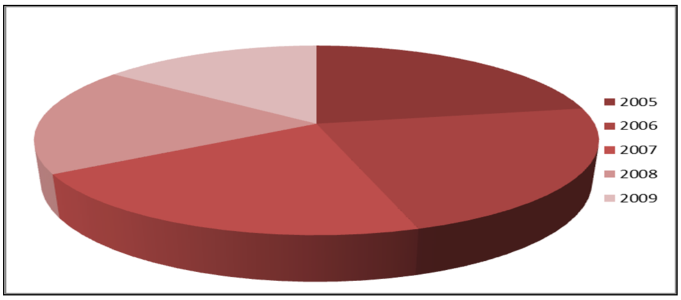


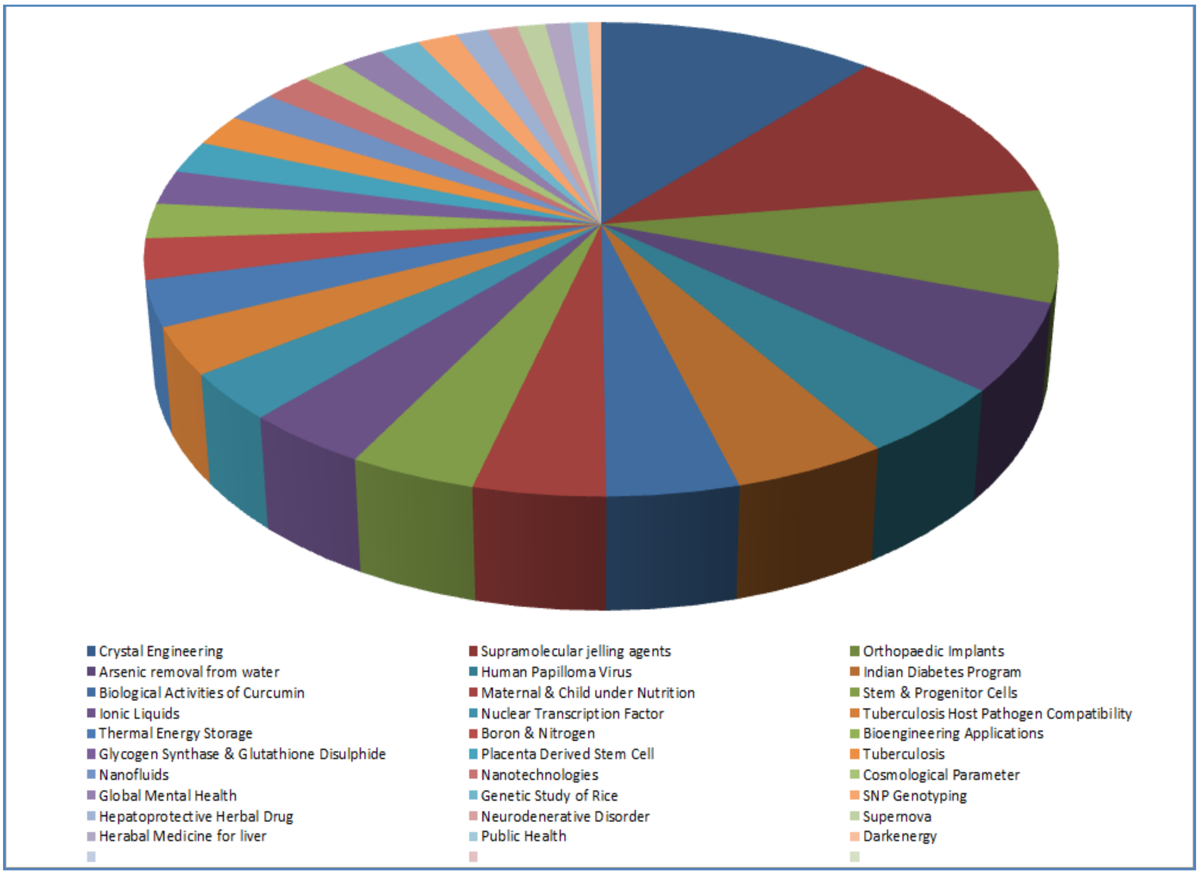
Figure 1. Year wise Citation Pattern to papers published during 2005-09
Source Data: WoS Expanded Online. Searched on March 2013

For computing the 'Citation' data, WoS (SCOPUS for calculating *h* index) has been used, as this has got a good coverage of Indian Papers being published in the area of Science & Technology including Medicine. More over the database is better structured and follow a standard formatting for the rendering of different data elements of particular field. The data base provides full support for 'Citation' as well as 'Co-citations Analyses'.

India's share of world papers and the relative number of citations to these papers received have both increased in recent years. However, while India is currently ranked seventh in terms of total output of papers within the group of Asian countries (SCOPUS data 2005-09), it remains tenth in terms of citation impact. The Impact Profiles for India's research publications show that while most of India's research is cited less frequently than the world average, India produces a significant volume of more frequently cited research. India published a total of 16 highly-cited papers (total 290- 700 Citations to each paper) in science and technology during the period of study (2005-09) as seen from the publications output data for 2005-09.

An analysis at the level of 'major discipline' indicated that two papers in the field of 'Particle Physics' stood at the **top with 3981 & 3679 Citations each**. Both the papers are review articles titled 'Review of particle physics' published in 'Physics Letters' & 'Journal of Physics G: Nuclear and Particle Physics' from Tata Institute of Fundamental Research, Mumbai (Bombay), with ≥ 100 co-authors from around the world, , indicating that the papers were the out- come of a metacentric study. The next one is a paper with maximum citation (543 citation) dealing in the area of 'Arsenic removal from water' . The other top 15 papers from different disciplines, receiving ≥ 200 Citations were from the field of Supra-molecular jelling agents, Orthopedic Implants, Arsenic removal from water, followed by Crystal Engineering, Human Papilloma Virus, Indian Diabetes Program, Biological Activities of *Curcumin*, Maternal & Child under Nutrition, Tuberculosis, Stem & Progenitor Cells, Ionic Liquids, Supra-molecular jelling agents, Crystal Engineering, Nuclear Transcription Factor, Tuberculosis Host Pathogen Compatibility, Thermal Energy Storage, Boron & Nitrogen and Biodiesel Production, Stem & Progenitor Cells, Ionic Liquids, Nuclear Transcription Factor, Boron & Nitrogen, Bioengineering Applications, Glycogen Synthase & Glutathione Disulphide and Placenta Derived Stem Cells.

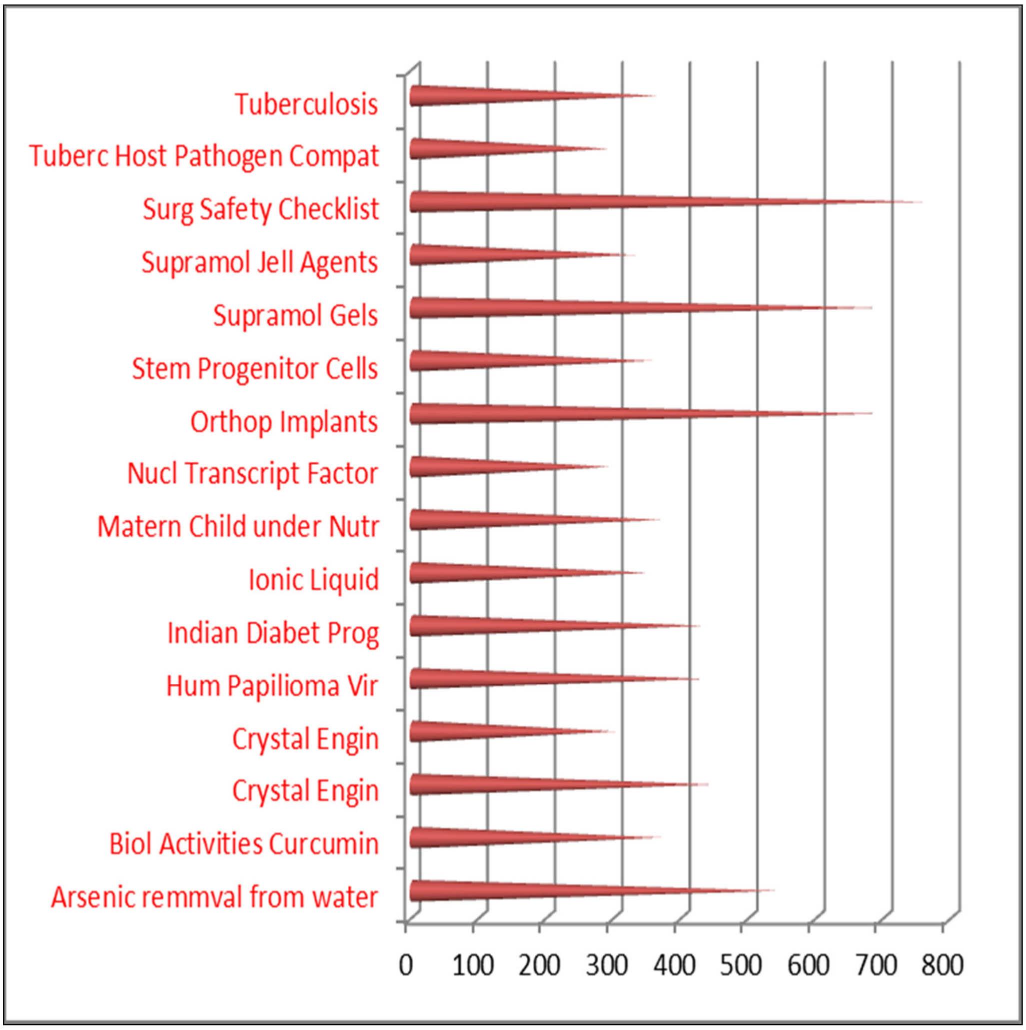
Subject Area Wise Citation Pattern



Source Data: WoS Expanded Online. Searched on March 2013

The next group of papers was of those having Papers with Citations in the range of 100 ≥ 249. The top most papers was from the field of 'Bioengineering Applications' followed by Glycogen Synthase & Glutathione Disulphide, Placenta Derived Stem Cell, Nano-fluids, Nanotechnologies, Chitosan, Global Mental Health, Genetic Study of Rice, SNP Genotyping, Effect of visual screening of cervix for Cancer, Hepato-protective Herbal Drug, Transgenic Crops, Neurodenerative Disorder and Tuberculosis.

Highly Cited Papers in the Broad Subject Areas of WoS.



Source Data: WoS Expanded Online. Searched on March 2013

An analysis was carried out to see the ranking of India, in terms of 'Citation' among Asian countries during 2005-09. During all these years Indian was among the top 10 countries occupying 3rd or 4th position throughout the period of study (***Source Data: SCOPUS Online. Searched on May 2013***).

Table 2.

2005

	Country	Total Papers	Citations	Cits / Paper	H index
1	Japan	108,184	1,270,580	11.49	602
2	China	154,940	968,120	6.22	353
3	South Korea	34,462	369,588	10.59	309
4	India	35,716	320,648	8.48	281
5	Taiwan	23,694	247,699	10.24	249
6	Hong Kong	10,634	157,518	14.27	268
7	Singapore	8,763	129,646	14.32	240
8	Thailand	4,242	57,974	13.28	156
9	Malaysia	2,971	20,325	6.66	116
10	Pakistan	2,506	15,811	5.92	101

2006

1	Japan	113,239	1,124,922	9.7	602
2	China	184,200	1,023,131	5.53	353
3	South Kore	40,313	348,298	8.51	309
4	India	42,572	328,389	7.31	281
5	Taiwan	26,788	240,166	8.77	249
6	Hong Kong	11,825	142,235	11.54	268
7	Singapore	10,380	123,565	11.49	240
8	Thailand	5,412	49,513	8.91	156
9	Malaysia	3,986	22,998	5.63	116
10	Pakistan	3,139	18,201	5.51	101

2007

1	China	208,313	1,035,319	4.94	353
2	Japan	109,110	953,575	8.52	602
3	South Korea	44,218	327,459	7.29	309
4	India	46,826	310,326	6.27	281
5	Taiwan	29,866	220,309	7.22	249
6	Hong Kong	11,943	133,631	10.77	268
7	Singapore	10,448	115,059	10.61	240
8	Thailand	6,188	52,135	8.26	156
9	Malaysia	4,700	24,356	5.09	116
10	Pakistan	3,791	21,157	5.35	101

2008

1	China	242,438	996,100	4.08	353
2	Japan	107,169	751,265	6.81	602
3	South Kore	47,170	293,390	6.11	309
4	India	51,904	266,414	4.85	281
5	Taiwan	32,132	194,491	5.91	249
6	Singapore	11,241	102,755	8.77	240
7	Hong Kong	11,742	102,684	8.43	268
8	Thailand	7,322	42,279	5.62	156
9	Malaysia	6,972	25,000	3.51	116
10	Pakistan	4,632	22,248	4.63	101

2009

1	China	284,372	823,957	2.87	353
2	Japan	107,025	531,006	4.8	602
3	South Korea	49,093	232,387	4.61	309
4	India	58,380	214,443	3.48	281
5	Taiwan	34,384	151,049	4.28	249
6	Singapore	11,702	81,989	6.73	240
7	Hong Kong	12,070	75,006	5.98	268
8	Thailand	7,599	34,797	4.4	156
9	Malaysia	10,262	28,389	2.72	116
10	Pakistan	5,664	17,543	3	101

Mapping of Indian Science

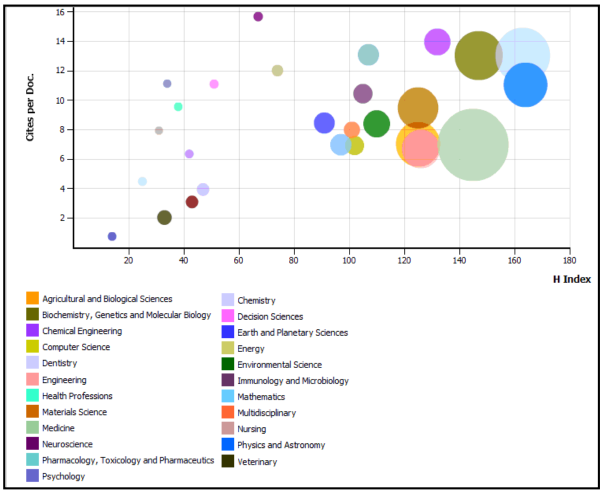
As we know that topics that are cognitively related to each other are positioned in each other's vicinity, and those not related or weakly related are positioned distant from each other. For strategic decision-makers in charge of national, institutional or R&D programs, the most valuable maps often are those based on co-occurring reference citations reveals. These co-citation maps reveal the cognitive structure of a scientific field, i.e., how the researchers themselves link different areas of knowledge.

The Map Generator software has been used to generate metrics to depict two types of Maps of Science & Technology, based upon national science indicators for 2 year periods between 2005 and 2009. These maps are intended to help reveal the existence of underlying scientific structures and plot science outputs and performance at a national level. The maps depict: Co-citation networks and Bubble charts of pattern of Citations and Co-citations. Both the maps has been generated for two time periods only; 2005-06 and 2008-09 to see if there is any change. Although, for this kind of study we need to do extended time periods for reaching to a conclusive results which is out of the scope of present project.

Bubble charts

Scientific output for India has been analyzed through customizable Bubble Charts for a richness of performance metrics. These charts also offer two levels of detail based on Scopus Classification's Science Areas (21 major fields) and Subject Categories (313 narrower thematic categories).Bubble charts features includes:

- Default view plots H index (on X axis) versus Citations / Papers (on Y axis), corresponding the bubble size to the field publication size.



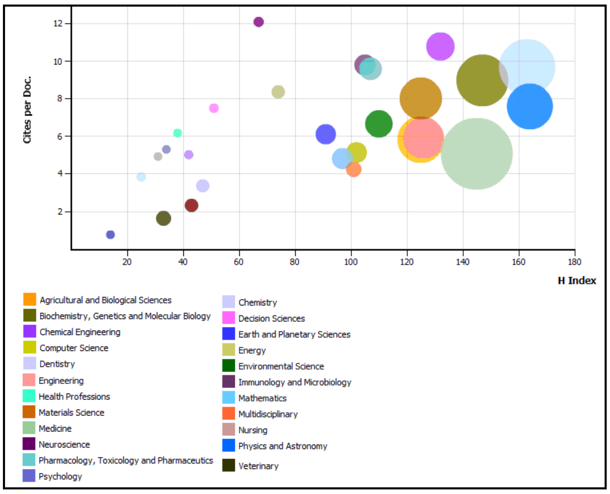


Figure 2. Bubble Chart of Citations / Paper and H index (2005-06)

Source Data: SCOPUS Online.

Bubble Chart of Citations / Paper and H index (2008-09)

Source Data: SCOPUS Online.

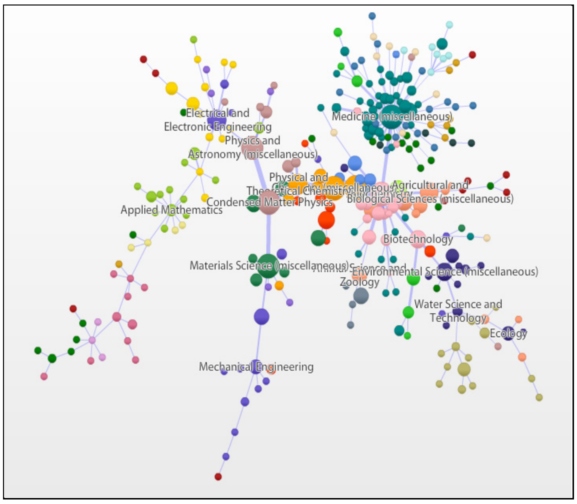
A close look at the 'Bubble Charts' of both the time periods (2005-06 & 2008-09) indicates that there are two main distinctly visible groups- Medicine in close 'contact' of Biological & Agricultural Sciences, Material Sciences and Engineering Sciences. This indicates a clear 'trend' of Publication / Citations of Indian Papers in the field of cutting edge technologies of Medicine like Nanotechnology .

The next group is of Mathematics linked with Physical Sciences & Astronomy, Biochemistry, Genetics and Molecular Biology; indicating towards impact of Indian research in the field of 'Bioinformatics' and Modern Biology.

Co-citation networks

National scientific structures were analyzed through two levels of detail for Co-citation Network Maps based on Scopus Classification's Science Areas (21 major fields) and Subject Categories (313 narrower thematic categories). The maps show the following features:

- Field/category size is depicted by the node size
- Relationship (similarity) intensity is displayed through thick / thin lines.



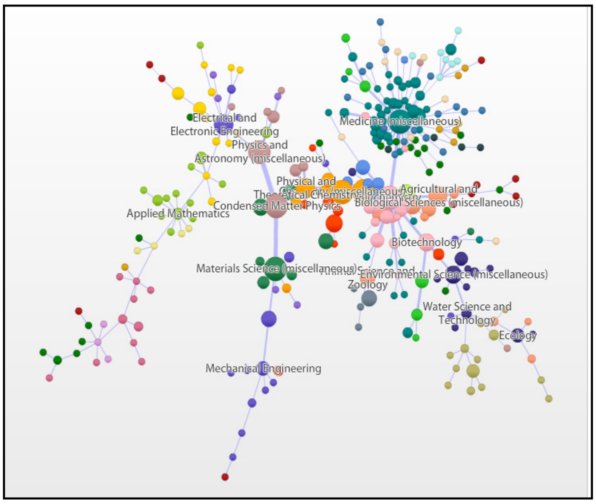
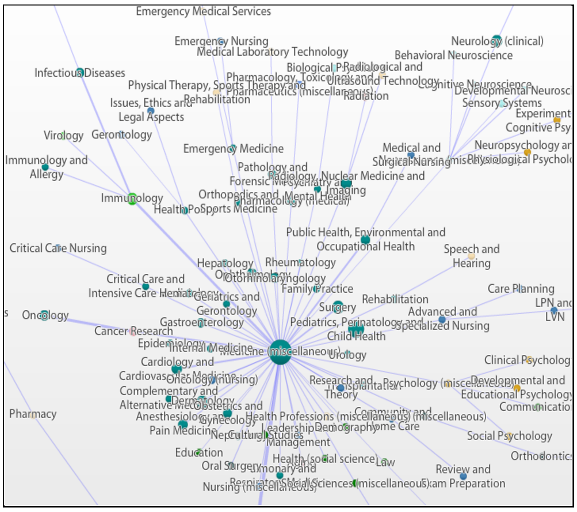


Figure 3. Co-Citation Map (with major subject disciplines) 2005-06

Co-Citation Map (with major subject disciplines) 2008-09



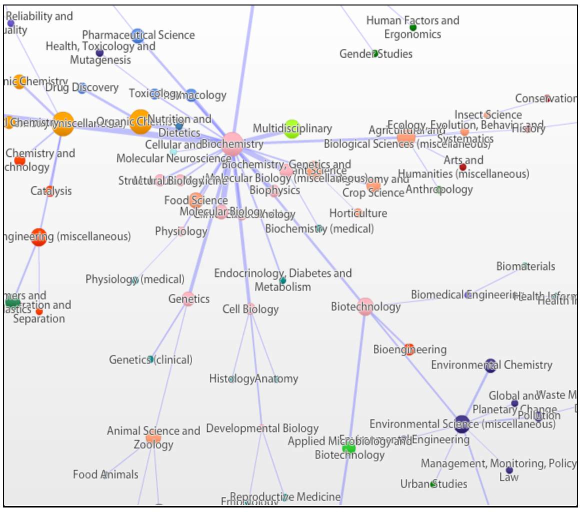


Figure 4. Medical Sciences **Biol Sci, Biochem & Environ Stud**

We should not forget, because co-citation-based maps reflect the connections between disciplines that thousands of researchers make each year, the linkages between disciplines can shift over time. From the 'Map' it is evident that the meta-discipline of biology no longer exists now. The majority of biology researchers during the period of study (2005-06 & 2008-09) focused their investigations on medical aspects of biology or on the chemical aspects of biology along with aspects of modern biology *eg.* Cell Biology, Developmental Biology, Embryology, Molecular Biology, Immunology or Reproductive Medicine. During the same years, the chemistry researchers drew more towards physical sciences and vice versa.

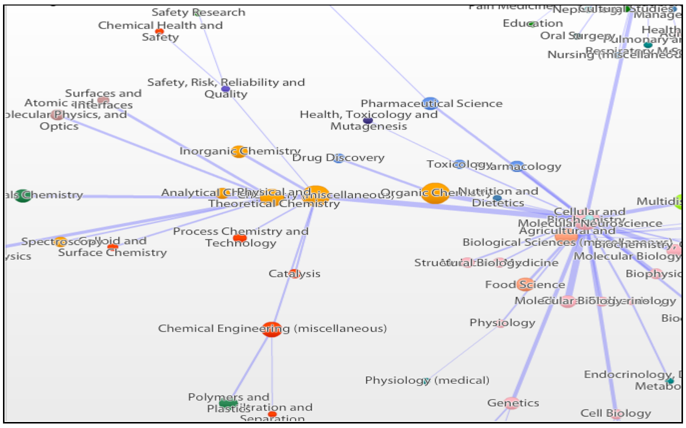


Figure 5. Chemical Sciences

Source Data: SCOPUS Online. Searched on May 2013

The Co-Citation Map for both the time periods (2005-06 & 2008-09), and Broad Areas along with Major discipline indicates that Medicine is very well established in terms of 'Citation Network' followed by 'Chemistry' and 'Material Science'. These maps shows that 'Medical Sciences' related research accounts for most of the science published during 2005-06 & 2008-09. In addition, 'Chemical Sciences research is linked to both biology and physics, but there no links between biology and physics.

The other subject fields with a little less impact at international level were as follows:

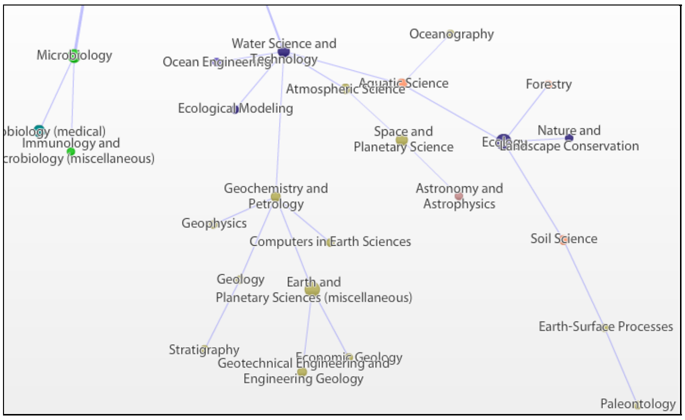


Figure 6. Earth & Atmospheric Sciences

Source Data: SCOPUS Online. Searched on May 2013

The field of Earth & Atmospheric Sciences, is an upcoming field for Indian researchers in terms of international visibility. None the less India has stepped into many new & very important frontiers as seen from the map like water science & technology, atmospheric science, geochemistry & petrology and Nature and conservation *etc.*

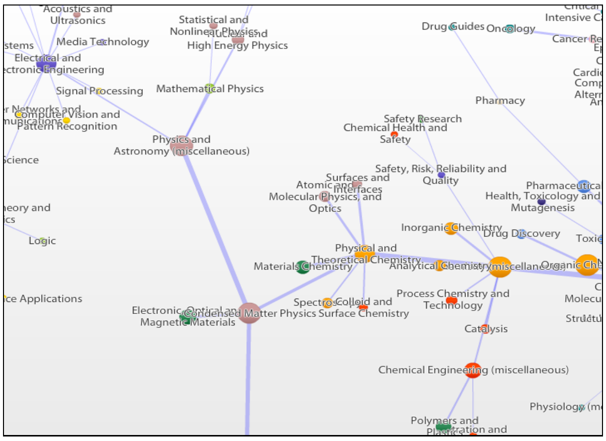
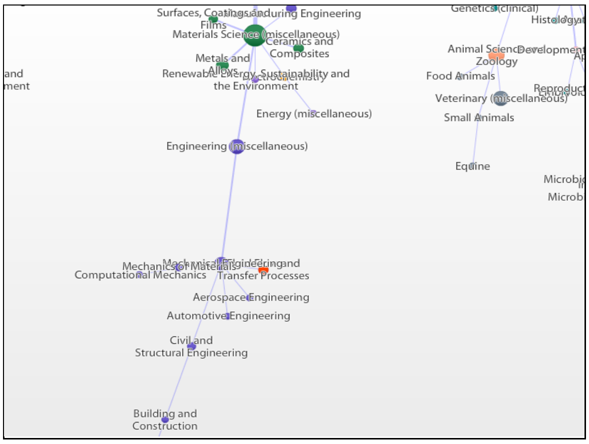


Figure 7. Engineering & Material Sciences



Physical Sciences & Astronomy

The other subject fields – Mathematics & Computer Sciences, Physical sciences & Engineering and Material sciences are yet to make their presence felt at international level in terms of Citations' to Indian papers from these areas.

While most of India's research is cited less frequently than world average it continues to improve. This growth suggests that India, along with other emerging economies, will become increasingly important to the global research community and that opportunities to collaborate with Indian researchers will increase.

References

Evidence from the latest BIS report indicates that international collaboration on articles boosts impact through citations and adds to the UK's position as a 'world-class' research nation. blogs.lse.ac.uk/impactofsocialsciences/2011/11/14/evidence-from-the-latest-bis-report/ Accessed on 20th Nov. 2012

Garg KC & Kumar Suresh (2010) Analysis of Citation Pattern of Indian Science Journals Indexed by SCIE. *Annals of Library and Information Studies* Vol. 57, 365-372

Leydesdorff L. (2007a). Visualization of the Citation Impact Environments of Scientific Journals: An online mapping exercise. *Journal of the American Society of Information Science and Technology*, 58(1), 207-222.

Moed H.F. (2005). Citation analysis in research evaluation. Dordrecht: Springer.

Osareh F.; McCain KW. (2008). Visualizing the structure of Iranian chemistry in Sci Search using author co-citation technique, *Collnet Journal of Scientometrics and Information Managemnt*, Vol 2(1):37-43.

Osareh, F. & McCain KW. (2008).The structure of Iranian chemistry research, 1990-2006: An author co-citation analysis. *Journal for Society of Information Science & Technology*, 59 (13): 2146-2155.

Schubert A & Braun T (1986). Relative indicators and relational charts for comparative assessment Of publication output and citation impact, Scientometrics ,9, 281-291.

Current status of Medical research across the Countries: India, China and Brazil

Divya Srivastava, Sandhya Diwakar and Ramesh Kundra

Division of Publication & Information, Indian Council of Medical Research, New Delhi, India
drdivya.srivastava@gmail.com, sandhyadiwakar@gmail.com, rameshkundra00@gmail.com

Rudimentary studies on aspects of medical research in India date back to 1911 with the establishment of 'Indian Medical Research Fund' which is now known as Indian Council of Medical Research. But, in the major fields of 'Medicine', such studies were started by scholars only during early 1970s with the establishment of research institutes under ICMR & Council of Scientific & Industrial Research and All India Institute of Medical Sciences, New Delhi, India. Science and Technology was not an immediate priority until 1961 due to domestic and political conditions in the country. We were then 11 years old since independence and our focus was on economic and social developments. Gradually, improvements were made in the field and now we have more 1200 to 1800 major research institution including medical colleges in the country, who have made significant contributions in the field during the last 8 to 10 years. Hence, we anticipate improvements in research output in terms of research papers in the field of Medicine from India. While examining the status and progress of Indian papers in the field of 'Medicine', this study also examines India's position vis-à-vis select developed and developing nations namely China & Brazil, in terms of its research output in the field of medicine over the years and major disciplines of research papers with a focus on India.

Medical research in India just like that in other nations, started as an offshoot of Biochemistry. Research in the area of Biochemistry in India can be traced to 1930s with publications related to "isolation of growth promoting factors in Bios" (Narayanan BT 1930) to "anticoagulant activity of fluorides, citrates and oxalates" to confirmation of classic work on "formation of Penicillin by P. notatum from lysine and β-hydroxy valine" (Deshmukh UD, et al 2001). This was followed by a big gap till thirty years later (in late 1960s), a group of dedicated biochemists like Profs. P. Sharma, G.P. Talwar, B.K. Bachhawat, M.C. Vaidya, D.P. Burma, C. Gopalan, A. Sreenivasan, L.K. Ramachandaran and K. Radhakrishnan plunged into active research with limited means.

An initial analysis of the data captured from SCOPUS for the period of 2001-2012, for papers appearing in the Broader area of medicine for India, China and Brazil has indicated an increasing trend over the years. Scopus has divided medical research into 47 further major disciplines.

Table 1. Total Papers in the field of Medicine

Year	*India*	*China*	*Brazil*
2001	3,644	3,153	3,213
2002	4,225	5,231	3,666
2003	5,045	7,139	4,155
2004	5,216	9,759	4,715
2005	5,960	14,137	5,489
2006	6,923	17,805	7,426
2007	7,523	20,979	8,229
2008	8,752	24,635	9,514
2009	9,680	27,493	10,416
2010	11,637	29,467	11,246
2011	13,228	32,007	12,086
2012	13,374	34,261	12,651

Papers from all the three countries were analysed further to compute data for the major disciplines being followed by these countries. The selection of these fields were entirely different from each other. During the period of study the 10 topmost disciplines were Pediatrics, Perinatology and Child Health, Surgery, Neurology (clinical), Dermatology, Public Health, Environmental and Occupational Health, Radiology, Nuclear Medicine and Imaging, Ophthalmology, Pathology and Forensic Medicine, Cardiology and Cardiovascular Medicine and Oncology. Whereas for China the focus was on Oncology research followed by Gastroenterology, Rehabilitation, Radiology, Nuclear Medicine and Imaging, Transplantation, Ophthalmology, Neurology (clinical), Complementary and Alternative Medicine, Surgery and Cardiology &Cardiovascular Medicine. Brazil has presented absolutely different areas. The topmost field was of Public Health, Environmental & Occupational Health followed by Surgery, Psychiatry & Mental Health, Cardiology & Cardiovascular Medicine, Infectious Diseases, Neurology (clinical), Obstetrics and Gynecology, Pediatrics, Perinatology & Child Health, Ophthalmology and Otorhinolaryngology.

Surprisingly, in the new and emerging areas *e.g.* Immunology and Allergy, Genetics (clinical), Transplantation, Critical Care and Intensive Care Medicine, Health Informatics India & Brazil are far behind than China. Chinese papers are maximum in the field of Oncology compared to India (Oncology-10th position) and Brazil (Oncology -19th position). With the growing number of Cancer cases this is of concern to policy makers. Some of the other emerging areas like Transplantation Medicine, Clinical Genetics, Radiology, Nuclear Medicine & Imaging, Clinical Neurology are also from the top 20 areas from China out of a total of 47 areas. Whereas for India the position is Transplantation Medicine (29th), Clinical Genetics (28th). In the field of Radiology, Nuclear Medicine & Imaging(6th), Clinical Neurology(3rd) India is ahead of China. In the case of Brazil, the topmost area is of Public Health, Environmental & Occupational Health. The area of Transplantation Medicine (31st), Clinical Genetics(28th), Radiology, Nuclear Medicine & Imaging (21st), are at much lower 'level'. In the areas of Clinical Neurology (6th) and Psychiatry & Mental Health (3rd) Brazil has produced much more papers as compared to China & India.

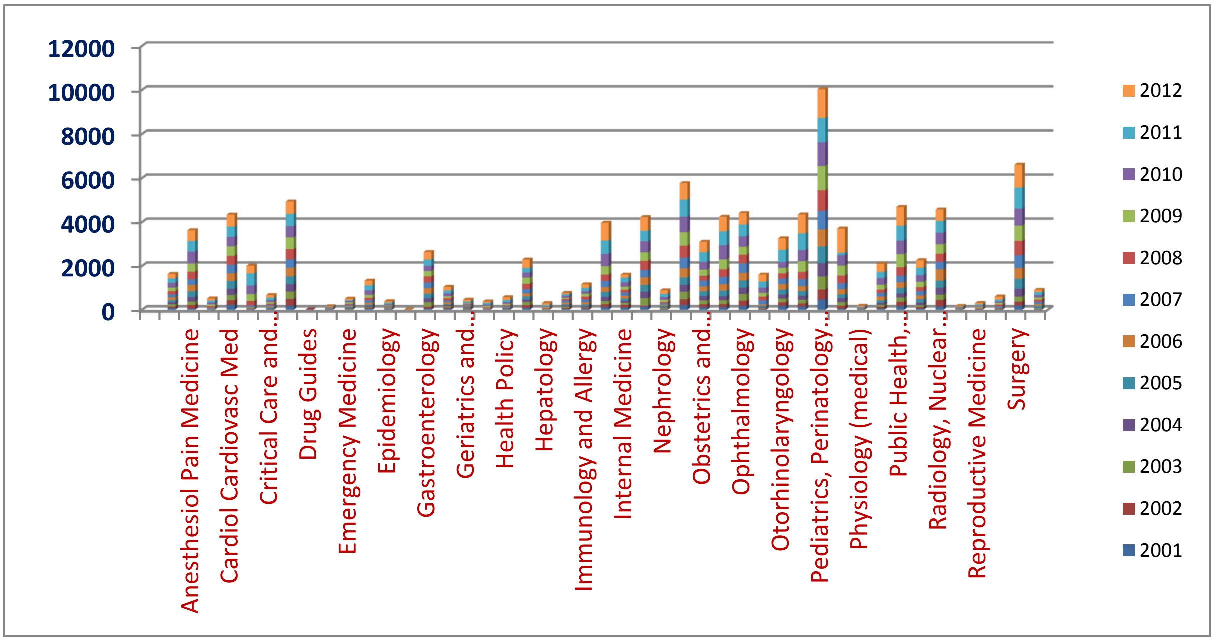


Figure 1. Total Papers-Major Disciplines: India

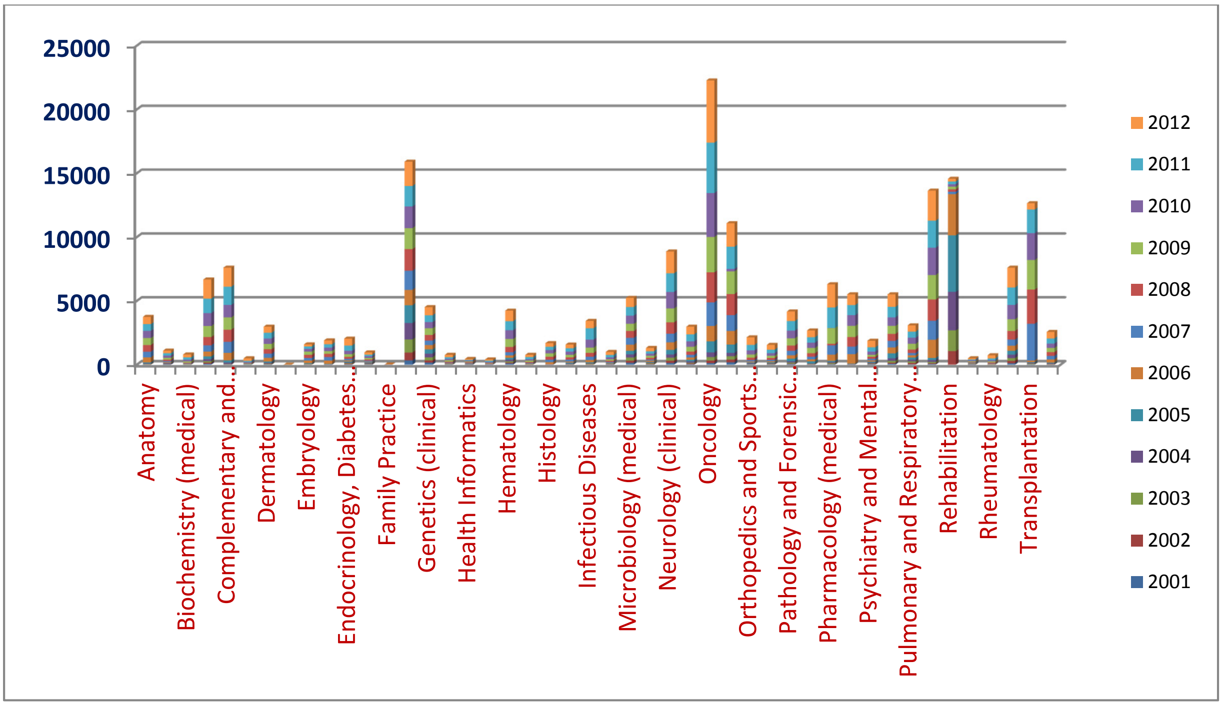


Figure 2. Total Papers-Major Disciplines: China

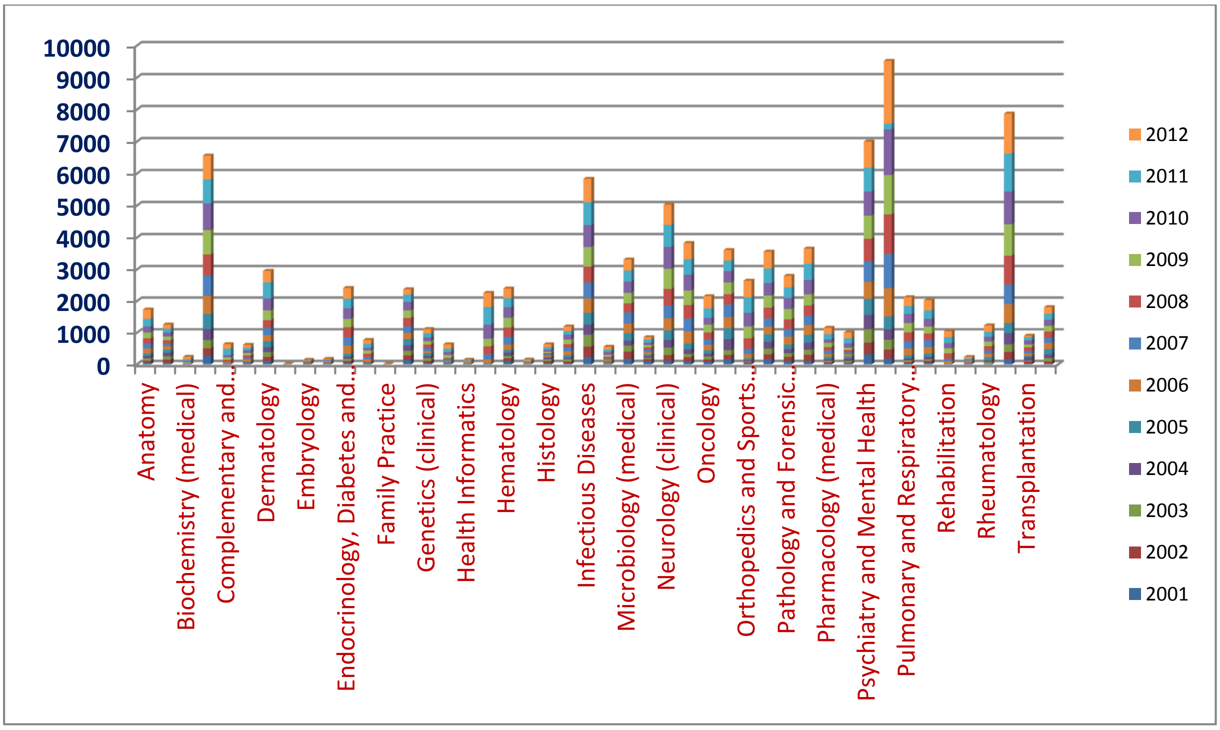


Figure 3. Total Papers-Major Disciplines: Brazil

To gain visibility in the area of medical research, India got chances to host the International Medical research Conferences across the country. The international community got exposed to strengths and weaknesses of the Indian medical research scenario. Ever since then, a gradual rise in clinical collaborations started, which continue to rise till date. In the meantime, over the period of last 20 to 25 years, scholars from the Indian Council of Medical Research, All India Institute of Medical Sciences, Department of Biotechnology, some of the institutions from Council of Scientific & Industrial Research and Department of Science & Technology worked hard to put Indian medical research on a serious path. However, while as the quality of work has considerably improved, the output is still low. One of the reasons for this could be that, most of the biological research in India is driven more by the "chemical perspective" but this may not be the only reason. Other reason is the lack of availability of manpower in the area. It would be worthwhile if more members of medical fraternity get drawn into the research area in addition to their basic interest in patient management.

Since its establishment, the ICMR has been making concerted efforts to address the health needs of the nation. Given its limited resources – human, financial and infrastructural the Council has discharged its national obligations through its network of 31 national institutes including Six regional medical research centres, over 100 field stations and a strong and vibrant extramural research in medical colleges and other institutes. The Indian Council of Medical Research (ICMR) has tried to address the issues relating to medical research and was also interested in capacity building at all levels for medical research, in partnership with international health organizations. The Council promotes Biomedical Research in the country through its several Human Resource Development programmes Diwakar S (2013).

Surprisingly, at the moment, we have 150 to 200 major scholars working in so many aspects of medical research while several others use medical research as a tool. The net outcome is a few great peer reviewed publications that could give great competition to big laboratories around the world e.g., 10 out of 60 molecular or protein structures of genomes of infectious diseases have been solved in India. This is a great feat indeed. One also needs to keep in mind that research infra-structure plays a great role in achieving goals. In the area of medical research of infectious diseases, P3 laboratories are being built or have been planned to be built at 10 to 12 places around the country so as to handle deadly organisms like Mycobacterium, Influenza virus, and HIV-AIDS. In order to run the show while as the number of key players remains more or less static, more key players need to jump into fray otherwise this could slow down the progress and hamper creation of a great bank in medical research research. A significant number of trained manpower leave India for their post doctoral studies, but very few of them return back in the area of medical research, unlike in other fields. The third component which results in lesser output is the funding position in the area of medical research. This has a bearing on the total spending on education and scientific research in India This matches the total budget of India for various activities for full one year under several departments including that of health and medical research. In spite of these impediments, collaborators proposals from abroad are growing and till we fine tune ourselves with several types of resources, it is advisable to prudently continue with multi-pronged strategy to give a push to the subject.

Health research is the key to a well-functioning and effective health sector in the country. Major scientific breakthroughs hold the promise for more effective prevention, management and treatment for an array of critical health problems. The research to be undertaken should be on country specific health problems essential for the formulation of sound policies and plans for field action. But new interventions and development of new health products (drugs, diagnostics and vaccines) are possible only when there is well defined funding, infrastructure and priority for health research. Medical research in the country needs to be focused on new therapeutic drugs/vaccines for tropical diseases, normally neglected by multinational pharmaceutical companies on account of their limited profitability potential. In addition, India is also witnessing the 'dual disease burden' with the non-communicable diseases like cardiovascular diseases, diabetes, cancers etc. threatening to overtake infections. In the Government sector, such research has been confined to the research institutions under the Indian Council of Medical Research, and other institutions funded by the Central/ State Governments.

We are witnessing a change in the field of medical research in India but at a slower pace due to paucity of scholars in the area. One very important filed that needs to grow at faster pace is in the area of disease resistance and susceptibility. There is a rapid need to improve infrastructure and add manpower in laboratories close to disease endemic areas so that local diseases and their remedies are better addressed.

References

Burnet FM, Fenner F. The Quarterly Review of Biology. 1951;26:220.

Burroughs Wellcome Fund, www.bwfund.org.

D Monte B, Talwar GP. J Neurochem. 1967;14:743.

Deshmukh UD, et al. J Postgrad Med. 2001;44:222

Gupta SK, Talwar GP. Indian J Exp Biol. 1980;18:1361.

Health Information of India 2004,Government of India, Central Bureau of Health Intelligence, DGHS, MOH & FW, Nirman Bhavan, New Delhi-11 website:www.cbhidghs.nic.in.

http://www.ias.ac.in/currsci/mar102001/599.pdf.

http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1522512.

Jha P, et al. Int J Lepr Other Mycobact Dis. 1971;39:14

Krishnaraj R, Talwar GP. Biochem J. 1972;128:57.

Mehra NK, et al. Lepr India. 1976;48:231.

Mitchison A. Nature. 1986;323:752.

Narayanan BT. Biochem J. 1930;24:6

National Commission on Macroeconomic and Health.

Rao RS. Indian J Med Res. 1952;40:341.

Report of the Working Group on Health Research for XII Five Year Plan (September 2011). Department of Health Research and Family Welfare, 1-2.

Verma RC, et al. Int J Lepr Other Mycobact Dis. 1971;39:20.

Visualizing the co-authorship relations in surgery discipline outputs among Iranian and Global cities

Farideh Osareh* and Ismael Mostafavi**

*Shahid Chamran University, School of Education and Psychology,
Department of Library and Information Science, Ahwaz, Iran
osareh.f@gmail.com

**PhD Student of Library and Information Science, Shahid Chamran University,
Department of Library and Information Science, Ahwaz, Iran
esmdoc@gmail.com

Abstract

In informetric studies, mapping of science uses a set of techniques that are useful to describe and analyze the growth of scientific performance and collaboration in scientific outputs. The purpose of this study is to overlay the network of relations among addresses in scientific publications onto the geographic map. In this study, scientific outputs of Iranian surgeons in Science Citation Index (SCI) via Web of Science (WoS) database, was extracted and mapped during 1990 to 2011. Application of study is that we overlay the organizational collaborations of scientists on the earth map. For example one can locate the organizations as nodes and produce links between them in Google Earth, however, does not allow us to show the global map at a single glance because of the globe format of the visualization. Totally, 1966 documents in Surgery were produced by 8789 authors, of those 3963 were Iranian and 4826 non-Iranian authors, which were extracted for the analysis. The geographic map based on both Iranian and Global co-authorship was drawn using Google Earth and Google Map. Highly co-authored articles were indicated on the map and the results showed that Iranian surgeons had more co-authorship links with European and North American countries compared to other continentals in the studied period.

Keywords: Informetrics, co-authorship, International co-authorship relations, Scientific outputs, Geography of surgery, Iranian surgery outputs, Mapping geography.

Introduction

Bibliometric measures have led to an increased interest among science and policy-makers for the identification of centers of excellence in scientific research (Danell, 2011). On the other hand, geographic regions and cities can be evaluated by describing new methods to analyze the geographic distribution of scientific production (Bormann et al. 2011). These methods allow to visualize the map of regions (and cities within them) that are characterized by co-authors and those authors who have published highly cited papers.

In informetric studies, mapping of science almost uses a set of techniques that are useful for describing and analyzing the growth of science and scientific collaboration in scientific output distributions. Many other disciplines, such as mathematics, statistics, computer science, and geography utilize mapping techniques. Geography is a field that uses an interdisciplinary approach to understand patterns of human activity and natural processes on the surface of the earth. Informetric professionals have been introduced with the use of these techniques to map the scientific outputs and study international scientific collaborations (Samll, 1999).

In this article, we first will review the relevant literature of surgery discipline that was published by collaboration of Iranian and Global surgeons and indexed in SCI via WoS database. Then the international collaboration network was mapped and subject to clusters based on Local Citation Score (LCS) and Global Citation Score (GCS) were identified. Co-authorship relations among Iran cities and global cities in surgery discipline was also visualized. We finally, sharing the unique insights gained from a global map by using Google map and Google earth.

In general, a map of science consists of a set of elements along with the relationships between those elements. These elements can be scientific fields or disciplines, journals, papers, or any other units that represent a partition of science. The characteristics that differentiate a map from a simple classification system are (a) the visualization of the elements, commonly represented by locating each element in a two-dimensional space, and (b) to explicit linking of pair elements by virtually the relationships between them. From the mapping perspective, classification is often thought of as a step along the way to creating a visual map, but is not equivalent with mapping if the relationships between the classes are not explicitly specified. Maps of science are commonly visualized as node-edge diagrams, similar to those used in network of science (Klavans and Boyack, 2009).

Research purpose

In order to analyze different aspects of scientific collaboration at the international level, a visual representation of surgery discipline was mapped. The main objective is to identify the international facet of research by following the flow of knowledge as expressed by the number of scientific publications, and then establishes the main geographical axes of output, showing the interrelationships of the domains. The intensity of these relations and how the different types of collaboration are reflected in terms of visibility also were studied. Thus, the methodology has two-fold application, allowing us to detect significant differences that help to characterize patterns of behavior of a geographical system of scientific output, along with the generation of representations that serve as interfaces for domain analysis and information retrieval. To reach the above goals, the following questions were raised during the years 1990-2011:

1. What are those countries and cities that their institutions have the most frequent co-authorship relations with Iran country and cities in Surgery scientific outputs?
2. Who are the most productive authors in surgery discipline?
3. Which Iranian cities have the most co-authorship relations with other Iranian cities in surgery scientific outputs?
4. How many clusters are there in the geographical map of Iranian Surgery scientific output?

Methodology & data gathering

Using Google Earth and Google Maps, and/or network visualization programs such as Pajek software, enables researcher to map the network of scientific collaborations based on the author's addresses. According to the city names, the global map can be drawn reliably on the basis of the available authors' addresses in the scientific outputs. For analysis of the Iranian surgery scientific outputs, we used scientometric method. Data was extracted using Science Citation Index (SCI) Expanded database via We of Science (WoS) on 18 February 2012. The results included 1966 records which have been published by at least one author affiliated of "Iran" address from 1990-2011.

In order to cover the available scientific literature, the search was performed in the topic field, which runs the search in titles, keywords and abstracts. Also advance search was run for searching the principal words used in related papers and "Iran" in the country field (CU). The format of scientific outputs studied in this research included articles, books, proceedings, book reviews, theses, letters. The data were analyzed using WoS analysis tool. Due to limitation of WoS, data were gathered in some 500 sets. All records in 500 sets Merged in one .txt file (data.txt).

Materials

The file (data.txt file) allows us to make a geographic mapp based on affiliated addresses and their relations using Google Earth. This input file is stored in the same folder as the programs cities1.exe and cities2.exe.

The two programs are to be run sequentially with an intermediate step. Cities1.exe is derived from isi.exe and first organizes the data into relational databases. It produces among other things a file named "cities.txt" which contains the city and country information (postcode if available) in standardized format. This file can be opened and then copy-and-pasted into the GPS encoder at http://www.gpsvisualizer.com/geocoder/. Then GPS Visualizer coded the city names.

The output of the geo-coding can be used as input into Cities2.exe after saving the file as a ".txt" file. The program prompts for the name of this file. It produces a number of output files in various formats (Leydesdorff, Persson, 2010):

1) Cities.kml and Cities2.kml can be read into Google Earth and/or uploaded to a website and then be read by Google Maps.
2) Network.kml contains only the network without the nodes.

Google Maps and Google Earth

The facility to read the files with extension .kml, into Google Maps provides us with many options to generate maps from the data by parsing and reformatting them into this rich markup language. However, the kml-language was primarily developed for Google Earth. Thus, the functionality in Google Maps is restricted to only a subset of tags. For example, one cannot scale the node sizes in Google Maps, but a user can do so by using the same file in Google Earth. Google Earth, however, does not allow us to show the global map at a single glance because of the globe format of the visualization, and has the noted disadvantage of only a single "satellite view" for the mapping. However, this image can be overlaid with street names and one can tilt the image. Google Earth is an online software that produces unique insights gained from a global map on the world. In Scientometrics we overlay the organizational collaborations of scientists on the earth map. Thus, the KML-language (one rich markup language) was primarily developed for Google Earth. For example one can locate the organizations as nodes and produce links between them by using the KML file format in Google Earth. Google Earth, however, does not allow us to show the global map at a single glance because of the globe format of the visualization. However, this image can be overlaid with street names and one can tilt the image (Leydesdorff, Persson, 2010).

The GPS Visualizer

GPS Visualizer at http://www.gpsvisualizer.com/geocoder/ allows us to input data either interactively or to read a file containing the required input information directly from one's disk. Running "Cities1.exe" produces a ".txt" file named "Cities.txt". This file includes the name of word cities that have collaboration in distribution of outputs. In the running of "Cities2.exe Software", for drawing the scientific collaboration maps, we need to degrees of "Longitude" and "Latitude" of cities. The GPS Visualizer noted above address can produce longitude and latitude degrees from the city names.

Institutional Collaboration

Focusing on collaboration among organizations, Mattessich and Monsey (1992) define collaboration as "a mutually beneficial and well-defined relationship entered into by two or more organizations to achieve common goals" (p. 7). They characterized the collaborative relationship as a durable and pervasive one, which aims to accomplish common goals (e.g., success and rewards) through a jointly structured and shared responsibility. Kagan (1991) also defines collaboration through organizational and inter-organizational structures where resources, power and authority are shared. People are brought together to achieve common goals, which could not be accomplished by a single individual or an independent organization. These two definitions are commonly used in the field of business and management, particularly in the management of joint ventures and strategic alliances among firms.

Literature Review

Glanzel and Schubert (2004) analyzed a co-authorship networks using bibliometric methods. In their study scientific collaboration was considered both at individual and national levels. Both literature data and original results witnessed a dramatic quantitative and structural change in the last decades of the 20th century. The changes, to great extent, can be attributed to the universal tendencies of globalization and the political restructuring of Europe. The standards and, particularly, the visibility of scientific research, as a rule, benefit from the ever increasing level of collaboration, but the profits do not come automatically. This fact underlines the necessity of a regular quantitative monitoring of inputs and outcomes, i.e., bibliometric surveys.

Bornmann et al. (2011)) studied the scientific collaboration among neuroscience authors field, using Scopus database in 2007. They, reported growing rate of collaboration at national and international level. Data was extracted using the search strategy "subject area (neuro) and pub year 2007 and doc type (ar)" in the advanced search field of Scopus. Circa 1% out of 40,082 documents (after ranking) equal to 407 records was selected for the sample of the study. Based on Scopus data, field-specific excellence can be identified and agglomerated in regions and cities where recently highly cited papers were published. First, unexpected encounters are more likely when two actors are in close vicinity of each other. Second, the need for face-to-face interaction when engaging in interactions comes at a cost, which increases as a function of travel time. Third, 'the rules of the game' that matter for scientific knowledge production (e.g., funding, labor market regimes, intellectual property right regimes, languages) are spatially differentiated and constrain interaction between institutional frameworks, in particular, between nation-states".

Bornmann and Leydesdorf (2011) mapped the top 10% highly cited papers published in 2008 in the three field: physics, chemistry, and psychology, using a citation window from this publication 2008 up to the date of harvesting data from the WoS for this research (February 2011). They mentioned the scientific mapping of excellent papers can complement the popular institutional rankings published so far.

Klavans and boyack (2011) describe two general approaches to creating document-level maps of science. To create a local map, one defines and directly maps a sample of data, such as all literature published in a set of information science journals. To create a global map of a research field, one maps "all of science" and then locates a literature sample within that full context. They provide a deductive argument that global mapping should create more accurate partitions of a research field than does local mapping, followed by practical reasons why this may not be so. The field of information science is then mapped at the document level using both local and global methods to provide a case illustration of the differences between the methods. Textual coherence is used to assess the accuracies of both maps. They find that document clusters in the global map have significantly higher coherence than do those in the local map, and that the global map provides unique insights into the field of information science that cannot be discerned from the local map. Specifically, we show that information science and computer science have a large interface and that computer science is the more progressive discipline at that interface. We also show that research communities in temporally linked threads have a much higher coherence than do isolated communities, and that this feature can be used to predict which threads will persist into a subsequent year. Methods that could increase the accuracy of both local and global maps in the future also are discussed.

Bornmann and Waltman (2011) by visualization methods (density maps) presented in this paper allow for an analysis revealing regions of excellence around the world using computer programs that are freely available. Based on Scopus and Web of Science data, field-specific and field-overlapping scientific excellence can be identified in broader regions where high quality papers (highly cited papers or papers published in Nature or Science) were published. They used a geographic information system to produce their density maps by Google Earth. They overlay map of authors in Europe having published highly cited biochemistry, genetics & molecular biology papers in 2007.

Leydesdorf and Persson (2010) by using the Google Earth, Google Maps, and network visualization programs Pajek, overlay the network of relations among addresses in scientific publications onto the geographic map. For mapping of science their data was extracted from ISI Web of Science and Scopus databases in library and information science journals. They used Google map, Google Earth, CiteSpace, the GPS visualizer, and Pajak Sotfware. Their study showed the scientific network relationship between cities of the world. They presented the collaboration networks on the Google Earth and Google Map.

Leydesdorff and Bornmann (in press) used a technique which is developed for patent information available online (at the US Patent and Trademark Office) useful for the generation of Google Maps. The overlays indicate both the quantity and quality of patents at the city level. This information is relevant for research questions in technology analysis, innovation studies and evolutionary economics, as well as economic geography. The resulting maps can also be relevant for technological innovation policies and R&D management, because the US market can be considered the leading market for patenting and patent competition. In addition to the maps, the routines provide quantitative data about the patents for statistical analysis. The cities on the map are colored according to the results of significance tests. The overlays are explored for the Netherlands as a "national system of innovations," and further elaborated in two cases of emerging technologies: "RNA interference" and "nanotechnology."

Results

In response to the research questions 1, from 5 continents of the world, only the cities of 44 countries have co-authorship links with Iranian cities in surgery discipline during 1990-2011. Of those, USA with 262 links, England with 51, and Canada with40 links ranked the top of the list respectively (Table 1).

Table 1: The countries that had collaboration with Iranian authors

	Country	No. of coll.		Country	No. of coll.
1	Iran	3963	11	Scotland	10
2	USA	262	12	Japan	9
3	England	51	13	Grenada	8
4	Canada	40	14	Spain	7
5	Germany	29	15	Austria, Belgium	5
6	Netherlands	23	16	China, Switzerland	4
7	France, Sweden	20	17	India, Jordan, Malaysia, Norway, Portugal, Saudi Arabia, South Korea	3
8	Italy	13	18	Brazil, Finland, Mexico, Poland, Qatar, Plestina	2
9	Australia, Turkey	Each 12	19	Argentina, Chile, Denmark, Egypt, Greece, Lebanon, Morocco, Singapore, Taiwan, Emirates, Wales	1

As was mentioned and also can be seen in table 1, the Iranian surgeons have had scientific collaboration with 3963 Iranian authors at national level, and 4826 global authors at international level Authors during1990 to 2011. Figure 1 shows a general view of Iranian cities co-authorship relations at the national and international level more clearly. As can be seen in this Figure 1, the most co-authorship links of Iranian cities in surgery outputs have been to European and North American countries.

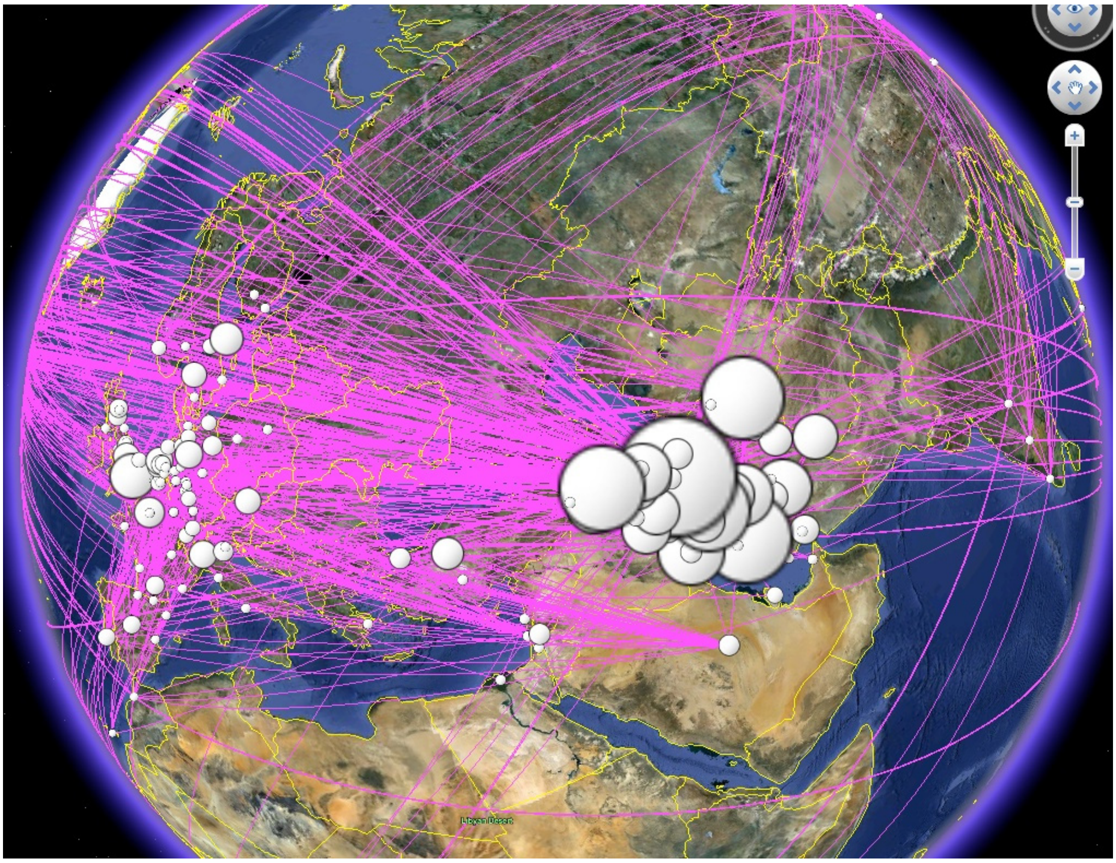


Figure 1: A general view of global cities that have co-authorship relations with Iranians cities in surgery discipline during 1990-2011

Totally Iranian surgeons have had co-authorship relations with 310 cities of the world. The top 10 Iranian and world cities are presented in table 2.

Table 2: The cities that had collaboration with Iranian authors

	City, Iran	Number of outputs		City, Country	Number of outputs
1	TEHRAN, IRAN	2342	1	BIRMINGHAM AL, USA	33
2	SHIRAZ, IRAN	378	2	LONDON, UK	26
3	TABRIZ, IRAN	237	3	PHILADELPHIA PA, USA	18
4	MASHHAD, IRAN	221	4	BOSTON MA, USA	16
5	ESFAHAN, IRAN	206	5	DALLAS TX, USA	14
6	AHWAZ, IRAN	74	6	LOS ANGELES CA, USA	14
7	YAZD, IRAN	56	7	MONTREAL, CANADA	14
8	KERMAN, IRAN	44	8	NEW ORLEANS LA, USA	12
9	ORUMIYEH, IRAN	39	9	INDIANAPOLIS IN, USA	10
10	RASHT, IRAN	39	10	STOCKHOLM, SWEDEN	9

In response to the second research question to identify the most productive Iranian authors in surgery discipline during 1990-2011; the most productive Iranian surgeons based on their co-authorship links are ranked and displayed in Table 3.

Table 3: Iranian most productive authors in surgery science

	Author	No. of outputs		Author	No. of outputs
1	Hashemi	47	7	Rajabi, MT; Tubbs, RS	20
2	Javadi, MA; Karimi, A	28	8	Marzban, M	19
3	Simforoosh, N	25	9	Basiri, A; Ghaemmaghami, F; Mehravaran, S; Najarian, S; Soheilian, M	18
4	Yazdani, S	24	10	Behtash, N; Kajbafzadeh, AM; Mandegar, MH; Moghimi, S; Nejat, F; Shoja, MM	17
5	Ahmadieh, H	23	11	Bayat, M; Falahatkar, S; Sheikhvatan, M	16
6	Karimian, F	21	12	Abbasi, K; Abbasi, SH; Davoodi, S; Zarchi, MK;	15

To find out which Iranian cities have the more co-authorship links in surgery discipline with the other Iranian cities? Figure 2 was drawn.

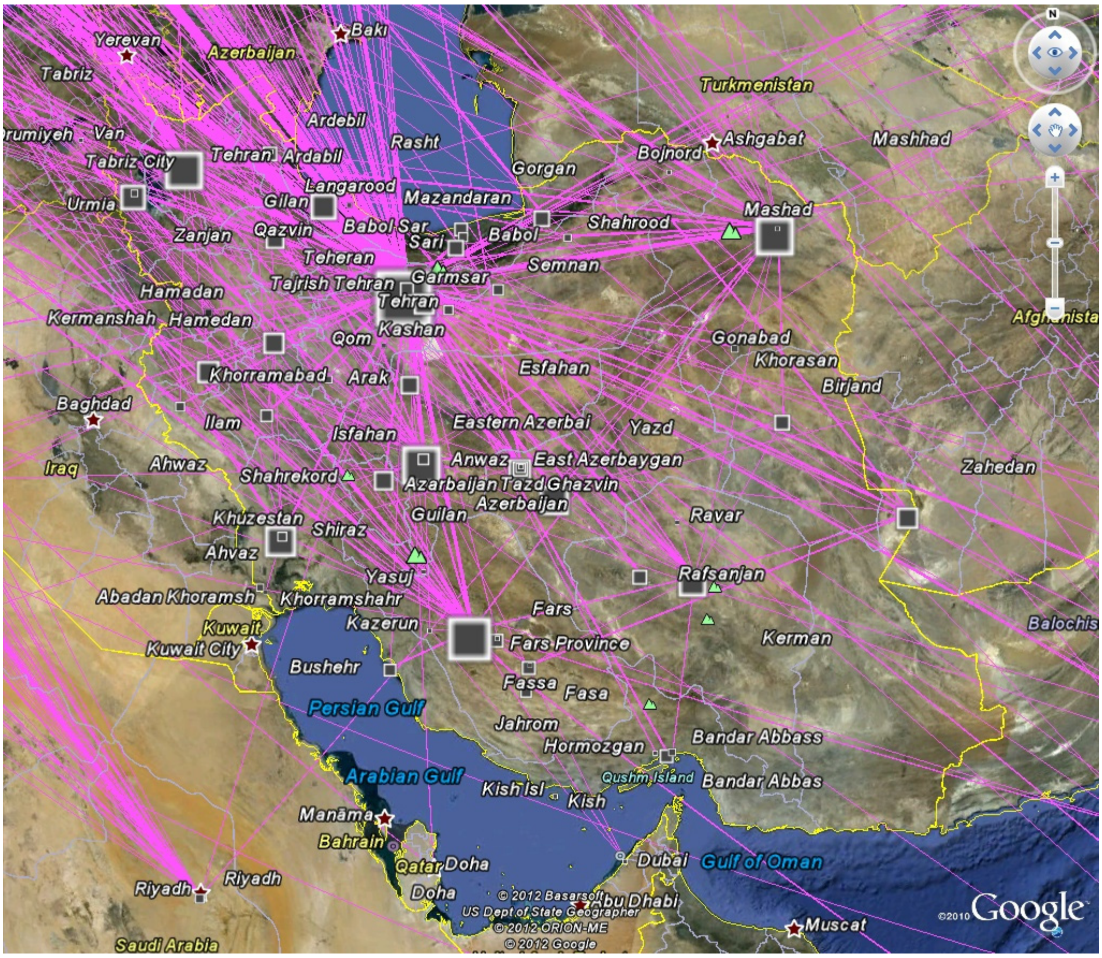


Figure 2: Iranian cities with co-authorship relations in Surgery discipline to other Iranian cities

As can be seen in this Figure, Tehran with 2342 co-authorship links ranked the first. Following Tehran, Shiraz (278) and Tabriz (237) ranked second and third respectively. The co-authorship relations of other Iranian big cities like Ahwaz, Isfahan, and Mashhad… are also displayed and can be seen on the Figure 2.

Co-authorship clusters among Iranian cities and global cities

One of the research questions is about the co-authorship clusters among Iranian cities and global cities on the geographical map of Iranian Surgery discipline during 1990-2011.

The result of the study showed 3 main co-authorship clusters among Iranian cities and other cities of the word. The biggest co-authorship cluster was formed among Iranian and European cities in surgery discipline during the studied period (Figure 3a).

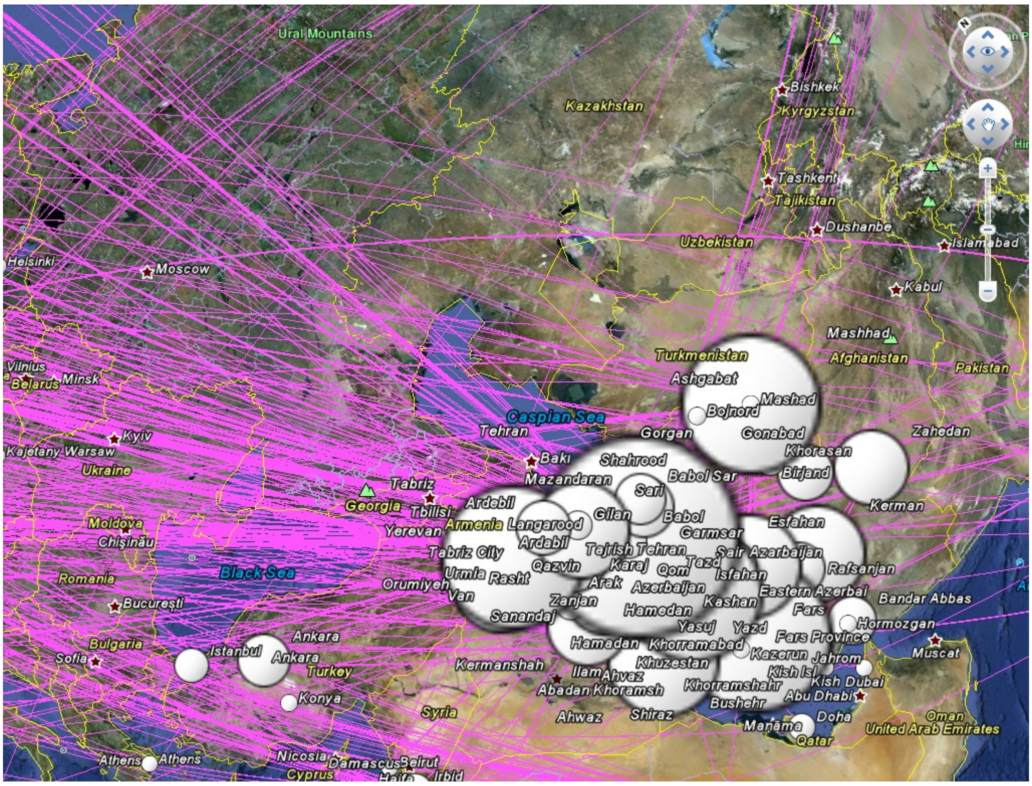


Figure 3a: Cluster 1: Iranian cities with co-authorship relations European cities in Surgery discipline during 1990-2011

Totally Iranian surgeons collaborated with their colleagues in 140 European cities, among them, cities like London, Paris, Berlin, Belgrade; Brussels…are clearly displayed in Figure 3a1. As can be seen in Figure 3a2 the countries' capital names are indicted by red stars.

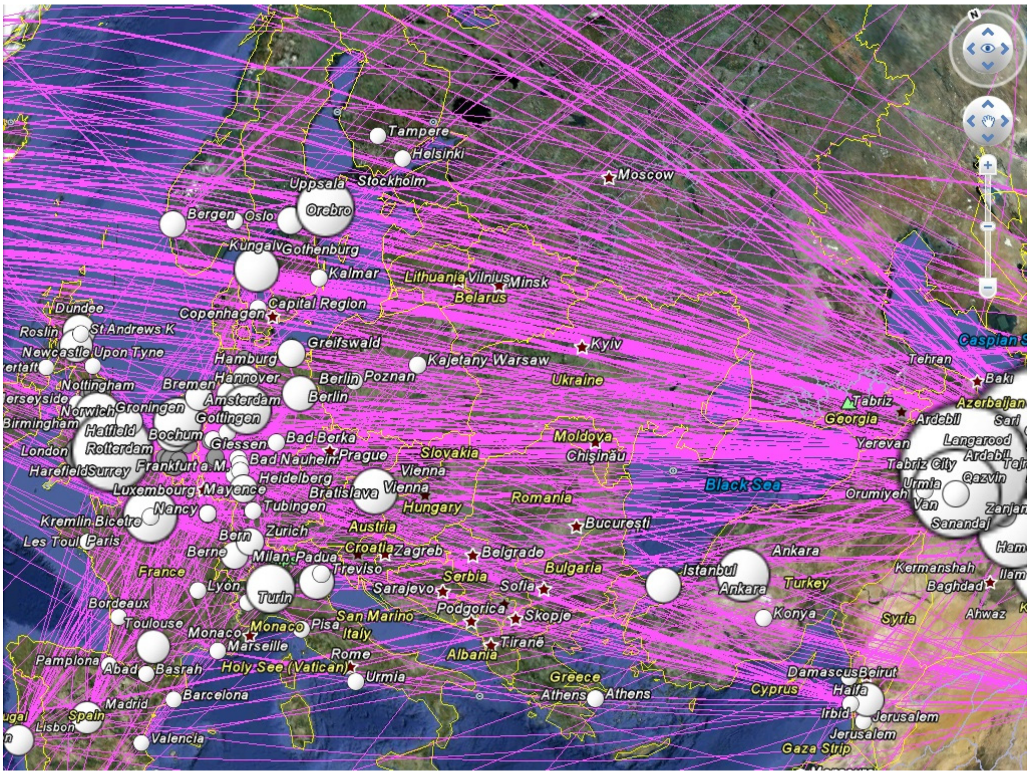


Figure 3a2: European cities that have co-authorship links with Iranian scientists in Surgery discipline during 1990-2011

The second biggest cluster in this study focuses on Iranian cities based on co-authorship relations to North American cities. As it is displayed on Figure 3b, most of Iranian co-authorship links to USA are to eastern, middle and southern cities such as New York city, Philadelphia, Pittsburg, Dallas, Kansas city, …. However, Iranian cities co-authorship links to the west of USA are few, and to some cities like Mexico city, Portland, Seattel…(Figure 3b).

Figure 3b1 also shows the co-atuhorship relations of Iranian cities to Canadian cities in surgery discipline during studied period. As it is displayed on this Figure, Canadian cities are strongly linked with Iranian cities in Eastern part of Canada, cities like Montreal, Toronto…but Canadian cities in west of the map (Figure 3b2) are linked with fewer links to Vancouver,…

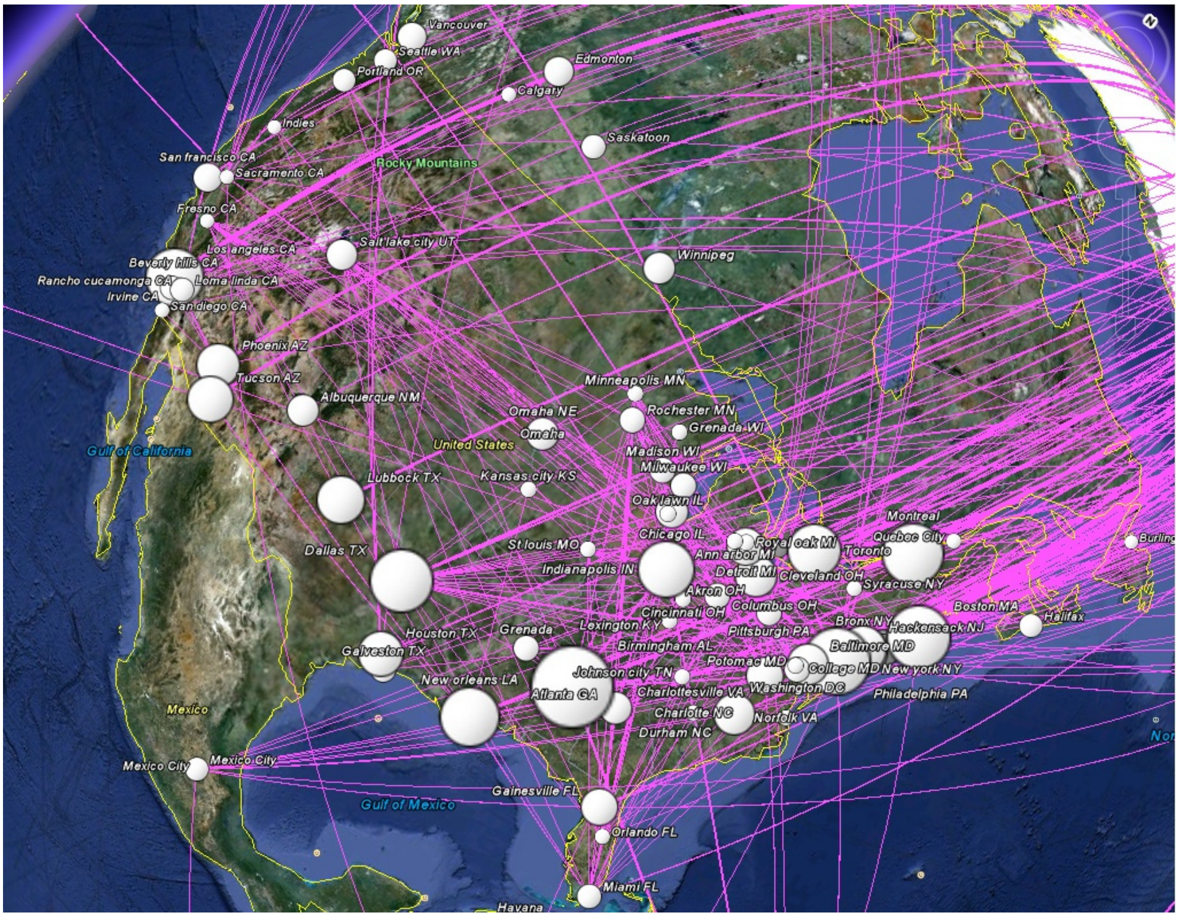


Figure 3b2: Iranian cities with co-authorship relations to North American cities in Surgery discipline during 1990-2011

Figure 3c shows a more clear view of the North American cities and the co-authorship relations with Iranian cities in surgery discipline during 1990-2011.

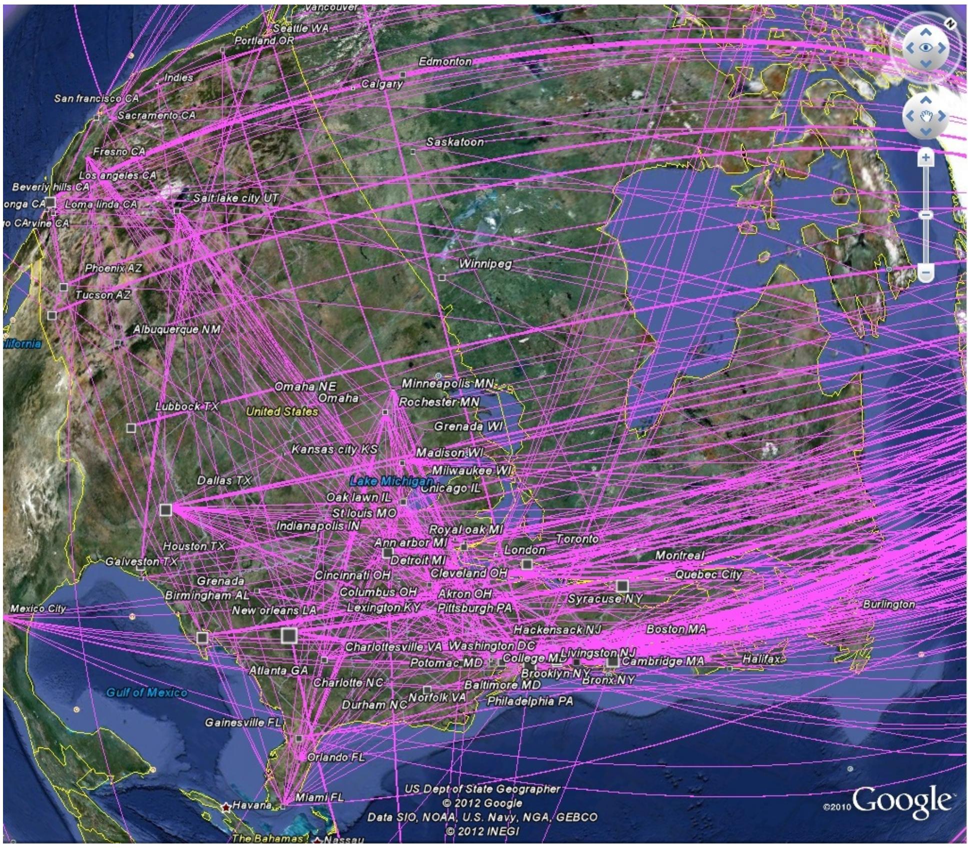


Figure 3c: Iranian cities with co-authorship relations to North American cities in Surgery discipline during 1990-2011

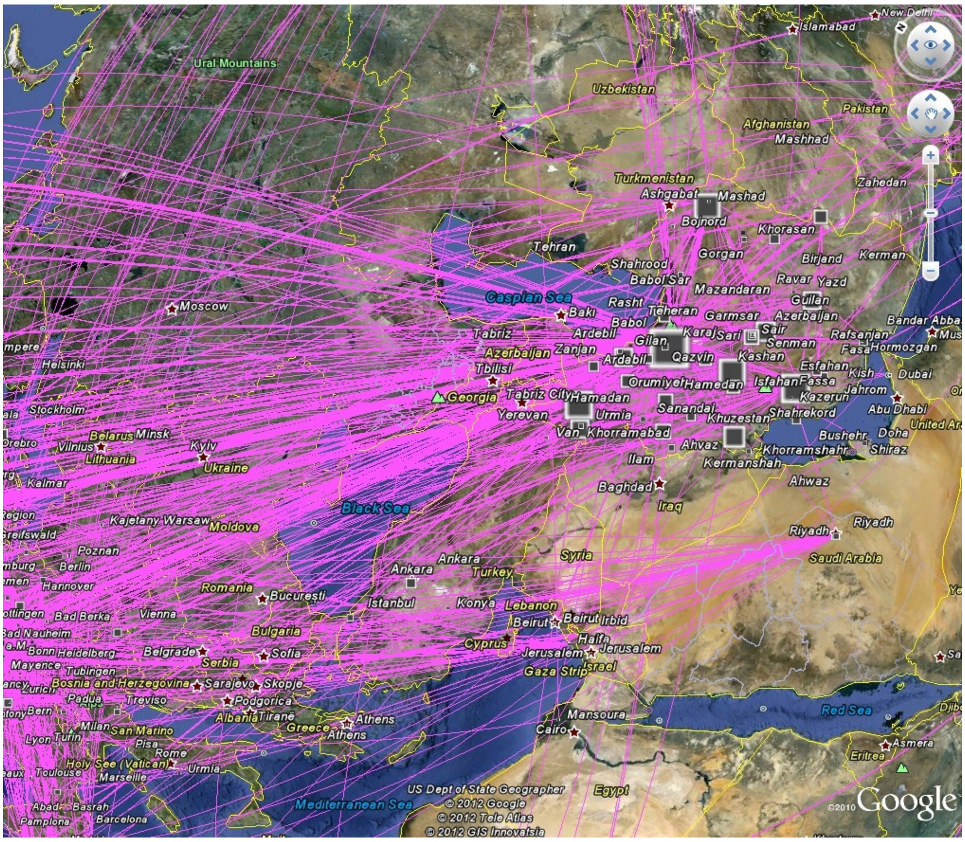


Figure 3d: Iranian cities with co-authorship relations to Global cities in Surgery discipline during 1990-2011

In general, despite, our Pointed out to the all clusters (Figures 3a-3c) of this article which displayed a global cluster based on co-authorship relations among Iranian cities with global cities; Iranian cities have co-authorship links with many other global cities distributed in all 5 continentals like: Moscow, Helshinki, Kyfv, Ukraine, Ankara, cyprus, syria, Iraq, Cairo, Riyadh…(Figure 3d).

Conclusion

The results of the study showed that in SCI Expanded database there were 1966 records in surgery discipline during 1990-2011 affiliated at least with one Iranian author. Totally, 1966 documents in Surgery were produced by collaboration of 8789 authors; of those 3963 were Iranian and 4826 non-Iranian authors. Google Earth was used to visualize data based on authors' addresses, in this research. The most Productive Iranian authors in this research were introduced. The results also showed that Iranian authors have had a vast scientific collaboration with cities in European countries, North American countries and the other countries of the world respectively. Iranian surgeons also have had strong co-authorship relations with their Iranian colleagues in Iranian cities at the national level.

References

Bornmann, L; Leydesdorff, L. Which Cities Produce More Excellent Papers Than Can Be Expected? A New Mapping Approach, Using Google Maps, Based on Statistical Significance Testing. *Journal of the American Society for Information Science and Technology*, 62(10):1954–1962. DOI: 10.1002/asi.21611.

Bornmann, L; Leydesdorff, L; Walch-Solimena, C; Ettl, C. (2011). Mapping excellence in the geography of science: An approach based on scopus data. *Journal of Informetrics*, 5: 537-546. Doi: 10.1016/j.joi.2011.05.005.

Bornmmann, L; Waltman, L. (2011). The detection of "hot regions" in the geography of science-A visualization approach by using density maps. *Journal of Informetrics*, 5: 547-553. Doi: 10.1016/j.joi.2011.04.006.

Danell, R. (2011). Can the quality of scientific work be predicted using information on the author's track record? *Journal of the American Society for Information Science and Technology*, 62(1), 50–60. doi:10.1002/asi.21454.

Glanzel, G; and Schubert, A. (2004). Analyzing scientific networks through co-authorship. *Handbook of Quantitative Science and Technology Research*, 257-276. H.F. Moed et al. (eds.), Netherlands: Kluwer Academic Publishers.

Kagan, S. L. (1991). *United we stand: Collaboration for child care and early education services*. New York: Teachers College Press.

Klavans, R. and Boyack, K. W. (2009). Toward a consensus map of science. *Journal of the American Society for Information Science and Technology*, 60: 455–476. doi: 10.1002/asi.20991.

Klavans, R. and Boyack, K. W. (2011), Using global mapping to create more accurate document-level maps of research fields. *Journal of the American Society for Information Science and Technology*, 62: 1–18. doi: 10.1002/asi.21444

Leydesdorff, L and Bornmann (in press). Mapping (USPTO) Patent Data using Overlays to Google Maps. *Journal of the American Society for Information Science and Technology*.

Leydesdorff, L and Persson, O. (2010). Mapping the Geography of Science: Distribution Patterns and Networks of Relations Among Cities and Institutes. *Journal of the American Society for Information Science and Technology*, 61(8):1622–1634.

Mattessich P. & Monsey, B. R. (1992). *Collaboration: What makes it work*. Saint Paul, MN: Amherst H. Wilder Foundation.

Rafols, I; Porter, A; Leydesdorff, L(2010). Science Overlay Maps: A New Tool for Research Policy and Library Management. *Journal of the American Society for Information Science and Technology*, 61(9):1871–1887. DOI: 10.1002/asi.21368.

Small, H. (1999). Visualizing science by citation mapping. *Journal of the American Society for Information Science* 50(9), 799-813.

Studying the status of knowledge management components in Petrochemical Companies (case study: South Pars Energy Economic Special Zone » Assalouyeh «)

Fatemeh Helaliyan Motlagh and Mohammad Hassanzadeh

fatemeh.helaliyan@modares.ac.ir, hasanzadeh@modares.ac.ir

Abstract

Purpose: today, knowledge is considered as human capability and organizational capital and its role is undeniable in organizational performance splendor and promotion. Hence, its management is getting increasingly importance. To this end, it is too vital to measure knowledge management status in oil, gas and petrochemical industry due to rapid technological advancement, merging many companies, high number of employees and the role of knowledge capital especially in South Pars Energy Economic Special Zone.

Methodology: this is descriptive survey. Its population consists of 370 employees of active companies in South Pars Energy Economic Special Zone. Relevant data is gathered by Proust et al. questionnaire and analyzed by deductive and descriptive statistics (t student and Freedman tests).

Results: studied petrochemical companies evaluated the status of knowledge management and such components as recognition, education, development, sharing, using and retaining knowledge as proper. Likewise, among knowledge management aspects, knowledge recognition has the best (3.25) and knowledge education (2.57) has the lowest status.

Keywords: knowledge management, petrochemical companies, current status, Probst model

Introduction

We are living in information and knowledge age. One day, “knowledge is power” was replaced by “knowledge share is power” in knowledge based age (Davenport et al., 198). Undoubtedly, this the age in which knowledge is considered as one of the greatest competitive advantages of organizations in global economy. In today world where goods production and service deliver is extensively knowledge – oriented, knowledge is a critical asset to acquire competitive advantage. In recent years, due to the emergence of a new stage of global economic system as “knowledge – oriented economy”, organizational knowledge is incrementally recognized as the main source of economy and organizational success is depended to its intellectual capitals rather than physical resources, capital and tangible assets (Walczak, 2005).

In this vein, one cannot deny the matchless role of knowledge is organizational performance splendor and promotion so that many organizations have allotted a huge volume of their operations to it. Knowledge is seen as human capability and a strategy for organization. Since any resource needs to be managed, knowledge also needs also management. Radical philosophy of knowledge management can be seen as a sustainable asset and organizations can achieve their goals by investing on it (Massa and Testa, 2009). Therefore, it is necessary that organizational run their knowledge effectively in addition to necessary agility and resilience (Bacerra – Fernandez, 2000).

Concerning above points, one can claim that the bottleneck in current organizations is not capital and manpower management; rather, it is employees’ knowledge management.

Knowledge management is not a new concept and it backs to the history of work. Wiig Prousak (2009) asserts that the main root of knowledge management backs to 3000 B. C. even though it was not called as knowledge management in its specialized format.

Davenport and Prousak (1998) define knowledge management as a process to flow knowledge among organizational members as a tool to achieve innovation in processes, products and services, effective decision making and organizational adaptability to dynamic environment and competitive market.

Of the most important reasons that have caused organizations to show their tendency to knowledge management, one can point out that knowledge management increases productivity and profitability, fosters cooperation, leads into creativity growth, encouragement and innovation, aids the establishment and acceleration of knowledge transfer from sender to receiver, enhances organizational ability to combat information inflation phenomenon, gathers and stocks employees' knowledge before their likely left and helps the organization not run out the scene by increasing the level of awareness on rivals' guidelines, products and performance (Aminppor, 2006).

Therefore, knowledge management is seen as an independent research scope in organizational studies and a process to acquire competitive advantage (Gloet & Terziovski, 2004). One of the most important domestic organizations is South Pars Energy Economic Special Zone as the most important gas field zone which involves half of oil and gas reservoirs.

Oil, gas and petrochemical industry is one in which the most important components are specialty and experience and its dominating principles can be written and coded very hardly. Organizational knowledge asset includes knowledge and learning in the mind of experts and managers taught during manufacturing processes which have short cycles and lost rapidly if a certain structure is not registered for them. The solution is to establish a comprehensive system called knowledge center.

Despite of rapid technological progresses and merging many companies, geographical dispersion, diversity of facilities, high number and employees and the importance of capital knowledge, knowledge management plays a vital role in oil, gas and petrochemical industry. Now, many global oil companies such Statoil Hydro, Shell, Exxon Mobil, Petronas and other companies are conducting knowledge management activities and have institutionalized knowledge management teams in their organizational structure (Nasr Esfahani, Taheri and Goli (2008). Any organization is ranked in a certain level based on its activities on knowledge management and this level indicates its current status in the field of knowledge management. Present study is conducted to answer two fundamental questions:

1. How is the status of knowledge management components in the petrochemical companies at South Pars Energy Economic Special Zone?
2. In which components do knowledge management components enjoy better status in the petrochemical companies at South Pars Energy Economic Special Zone?

Present study aims at investigating the status of knowledge management components in the petrochemical companies at South Pars Energy Economic Special Zone (Assalouyeh).

By realizing the aims and results of the research, one can help petrochemical companies to employ committed workforce, to preventing wasting efforts and resources, to provide new products, to execute knowledge management successfully and so on.

Studying relevant literature on knowledge management indicates that there are paramount models in this regard. There is no main difference between them. The only difference is in the name and number of the steps of this process. Precise investigation of each model indicates that their steps are too similar and some authors have only considered this process more completely. By combining and aggregating similar steps in such models, one can observe knowledge management process fully in the building blocks of knowledge management by Proust, Rob and Wermhardt. Therefore, research conceptual model (graph 1) is expounded as below.

- In terms of "knowledge recognition", petrochemical companies at South Pars Energy Economic Special Zone are in proper status.
- In terms of "knowledge learning", petrochemical companies at South Pars Energy Economic Special Zone are in proper status.
- In terms of "knowledge development", petrochemical companies at South Pars Energy Economic Special Zone are in proper status.
- In terms of "knowledge sharing", petrochemical companies at South Pars Energy Economic Special Zone are in proper status.
- In terms of "knowledge utilization", petrochemical companies at South Pars Energy Economic Special Zone are in proper status.
- In terms of "knowledge retain", petrochemical companies at South Pars Energy Economic Special Zone are in proper status.

Research conceptual model

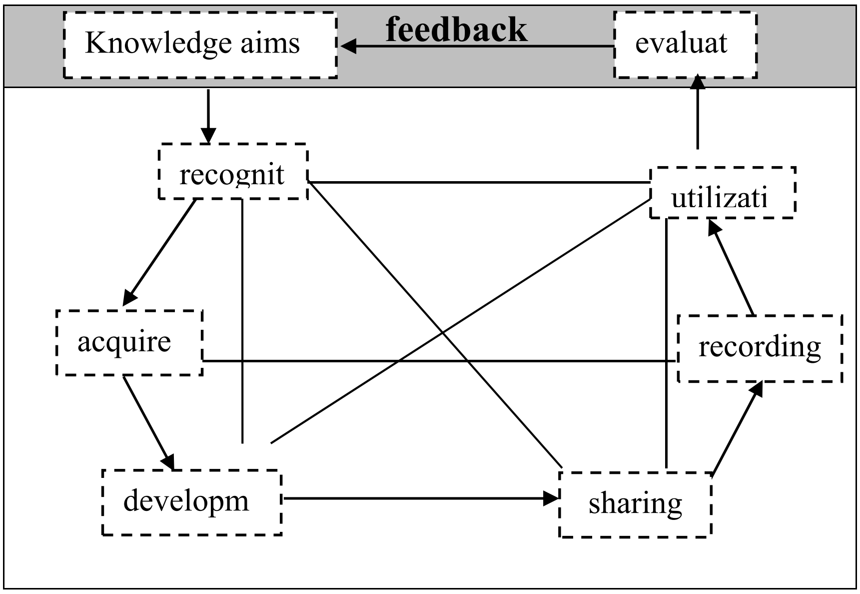


Figure 1. (Probst, Raub & Rombardt,2000)

Problem description and research importance

Today, equipping with updated information and knowledge is considered as a stable situation to survive individually and socially and the competition capability in the market depends on acquiring and developing individual and organization knowledge so that knowledge is considered as the main part of capital. In companies, knowledge management is seen as an individual need and a radical initiative to enter global competition and facing with business challenges (Miguel, 2007).

In recent years, knowledge management has become a critical issue. Scientific and trading communities believe that organizations can keep their long term superiorities in competitive arenas by their knowledge power. In their studies, authors have found that in contrary to other kinds of management, knowledge management is not temporary; rather it has permanent effects. Organizational competitive conditions are changing and complicating rapidly so that the velocity of changes in organizations is much more than their responsiveness and adaptability. Knowledge constant changes have caused new imbalanced situations for companies. In this way, only those organizations can survive that are able to retain their competitive advantage. According to practitioners, this era of keeping competitive advantage and organizational survival is feasible by the aids of knowledge management so that one can constantly generate new knowledge in organizations (Bakhtiari, 2009).

Concerning 20-year outlook document and its emphasis on creating and developing knowledge – based society and recognizing the status of knowledge management, it is obvious that knowledge management should be considered as an organizational necessity. The main problem in present study is that "how is the status of knowledge management components in petrochemical companies at South Pars Energy Economic Special Zone?

If such problem is resolved and the status of each knowledge management component is determined, then it would lead into advantages for companies as below:

Increase in cooperation; productivity improvement, determining organizational weak and strength points, correct evaluation, proper budgeting, facilitating knowledge sharing among employees and product increase (Latifi, 2007).

The reverse situation is also true namely the problems and disadvantages as below:

Lack of proper assessment tools, lack of proper budgeting, lack of signs of invention and initiative, absence of prioritizing the usage of knowledge types, non-observing external knowledge, non-attracting new knowledge – based groups (Keyvani, 2011).

Research background

Today, paramount researches are conducted on knowledge management scope which shows the importance of this discussion in knowledge management.

Yujiang X. Yuan Li (2009) conducted a study to investigate knowledge management and innovation in organization through a survey and a questionnaire by random sampling method in 127 German companies. Knowledge sharing, knowledge generation, performance and innovation in organization were considered as the factors of knowledge management while new operational ideas, new ways to perform the job, know – how techniques, new production and marketing skills were considered as assessment indices. The author concluded that learning organizations play a vital role in knowledge generation process, knowledge management impacts directly on organizational performance and innovation and joint investment facilitates more knowledge generation and sharing than contractual relations.

Kamali Tabrizi, Fariborz and Gholam Husseinzade, Zohreh (2011) conducted a study on measuring the readiness of Oil Research Center to execute knowledge management. Relevant information is gathered through a 51 – item structured questionnaire. Its population consists of 130 authors in this Center. This research measured the readiness of the center in two terms: 1. Organizational culture aspect, IT supports, organizational structure and their related process; and 2. Employees' beliefs in terms of profitability and easiness of executing knowledge management process. Authors concluded that concerning the first aspect, the center lacks readiness to define proper processes and while it is ready for the second aspect.

Methodology

In terms of type, this is applied study while it is a survey descriptive one in terms of data collection method. Total research population consists of 12,188 employees of petrochemical companies at South Pars Energy Economic Special Zone. Layer sampling method is used to conduct present study. Since it was impossible to conduct the research in total population, Morgan Table is used to determine sample size. On this basis, sample size is estimated as 370. Noteworthy, of 370 sent questionnaires, 317 were completed and returned. Therefore, the results are from 317 completed questionnaires. To gather data, knowledge management measurement questionnaire by Golvani (2008) is used. The only difference is that after mitigating the components and items and adding some questions, the questionnaire was divided into two parts. The first part was on demographical variables (age, gender, educational degree, location, job type and job record) while the second part consists of 27 five – item questions based on Likert 6-point scale based on Proust's building blocks of knowledge management. To compute reliability value, Chronbach's alpha value is used (a = 89%). Its figure shows the reliability of the questionnaire. Finally, to analyze data, descriptive statistical (relative frequency, average, standard deviation) and deductive (T-Student and Freedman) indicators are used.

Findings

Extracted information from gathered data was analyzed by descriptive and deductive techniques explained below separately:

Descriptive findings: this part describes respondents' general traits statistically. Gathered data is present study show that 95.6% of respondents were male. The most frequency of respondents was 24 – 35 years (74.1%) and the lowest frequency was 46 – 55 years (1.9%). Concerning the educations, research findings suggest that the highest one (55.5%) is bachelor while the lowest one (0.6%) was doctoral. In terms of job record, research findings indicates that 28.7% of respondents' job record is less than 5 years while 36% had 5 – 10T 25.2 had 10 – 15, 3.8% had 15 – 20 and 1.6% had job over 20 years of job records.

Table 1: the frequency distribution of employees in terms of gender, age, educations and job records

Gender	Percent	Age	Percent	Educations	Percent	Job records	Percent
Male	95.6	24 – 35	74.1	Diploma	12.6	Under 5	28.7
Female	4.1	36 – 45	16.7	Associate of arts	14.8	5 – 10	36
-	-	46 – 55	1.9	Bachelor	55.5	10 – 15	25.2
-	-	-	-	Masters	12.9	15 – 20	3.8
-	-	-	-	Doctoral	0.6	Over 20	1.6

Deductive findings

Here, we examine hypotheses by T test in SPSS software package. Then, by using Freedman test, six aspects of knowledge management are tested in petrochemical companies at South Pars Energy Economic Special Zone. After studying each hypothesis by using T Test in significant level of 0.05 and standard distance of 0.09, one can decide as below that which knowledge management components enjoys proper status.

In the case that significant level of one aspect is lower than 0.05, it knowledge management is not proper and null hypothesis is supported.

In the case that significant level of one aspect is greater than 0.05, it knowledge management is proper and H1 is supported.

Below, we test research hypotheses.

Hypotheses1. In terms of "knowledge recognition", petrochemical companies at South Pars Energy Economic Special Zone are in proper status.

Table 2: the results of T Student test to examine knowledge recognition

Components	Average	Standard deviation	t	Test result of H0 (OR H1)
Prioritization of organizational knowledge management	3.11	0.979	56.542	Proper
Organizational awareness of its knowledge weaknesses	3.08	0.875	62.553	Proper
Awareness of generated knowledge in organization	2.80	0.948	52.681	Proper
Finding capability for needed knowledge	3.49	0.833	74.647	Proper
Awareness of needed knowledge	3.80	0.795	85.169	Proper
Recognized value resources	3.32	0.839	70.355	Proper
Awareness of colleagues' knowledge	3.18	0.858	65.967	Proper
Recognizing an individual with the best answer to your questions	3.27	0.913	63.890	Proper

In table 2, descriptive statistics include average, standard deviation and t test results to show knowledge recognition and its indicators. Concerning above table, the highest average belongs to components 5 (3.80) while the lowest one is components 3 (2.80). With regard to achieved results, it is determined that knowledge management and its indicators are evaluated as proper by personnel of petrochemical companies. As a result, these companies are in proper situation in terms of knowledge recognition. So, Hypotheses 1 is supported.

Hypotheses2. In terms of "knowledge learning", petrochemical companies at South Pars Energy Economic Special Zone are in proper status.

Table 3: T Student test results for knowledge learning

Components	Average	Standard deviation	T	Test result of H0 (OR H1)
Using external and internal advisors by organization	2.68	1.189	40.148	Proper
Acknowledging employees for learning new knowledge	2.46	1.095	40.063	proper

As seen in table 3, the highest average relates to item 1 (2.68) while the lowest one is item 2 (2.46). As a result, t test shows the status of knowledge learning indicators with high confidence3 (sig = 0.000<0.05). It means that employees of petrochemical companies assess the status of knowledge learning as proper. Thus, Hypotheses 2 is supported.

Hypotheses3. In terms of "knowledge development", petrochemical companies at South Pars Energy Economic Special Zone are in proper status.

Table 4: T Student test results for knowledge development

Components	Average	Standard deviation	T	Test result of H0 (OR H1)
Increase in knowledge and experiences due to the operations in organization	3.62	0.965	66.868	Proper
A mission out of organization	1.79	1.130	28.276	Proper
The rate of holding training courses, seminars, …	2.54	1.135	39.851	Proper

Table 4 indicates that the highest rank relates to item 1 (3.62) while the lowest one is item 2 (1.79). Since significance level is 0.000 and lower than 0.05, Hypotheses 3 is supported.

Hypotheses4. In terms of "knowledge sharing", petrochemical companies at South Pars Energy Economic Special Zone are in proper status.

Table 5: T Student test results for knowledge sharing

Components	Average	Standard deviation	T	Test result of H0 (OR H1)
Employees'exchanges with each other	2.59	1.047	44.089	Proper
Collective works	3.30	1.047	56.102	Proper
Spent time to exchange information	3.13	0.897	62.155	Proper
The impactby employees' experiences and knowledge on organizational promotion	3.62	1.066	60.456	Proper

In table 5, t test results are shown for knowledge sharing and its indicators. Results indicate that the highest average relates to item 4 (3.62) while the lowest one is item 1 (2.59). To this end, null hypothesis on knowledge sharing is refused. It shows that petrochemical companies are in good status of knowledge sharing. Thus Hypotheses 4 is supported.

Hypotheses5. In terms of "knowledge utilization", petrochemical companies at South Pars Energy Economic Special Zone are in proper status.

Table 6: T Student test results for knowledge utilization

Components	Average	Standard deviation	T	Test result of H0 (OR H1)
Acknowledging employees for utilizing new knowledge	2.54	1.132	39.950	Proper
The rate of utilizing accessible knowledge	3.33	0.928	63.882	Proper

According to table 6, the highest average relates to item 2 (3.33) while the lowest one is item 1 (2.54). To this end, null hypothesis on knowledge utilization is refused. It shows that petrochemical companies are in good status of knowledge sharing. Thus, Hypotheses 5 is supported.

Hypotheses6. In terms of "knowledge retain", petrochemical companies at South Pars Energy Economic Special Zone are in proper status.

Table 7: T Student test results for knowledge retain

Components	Average	Standard deviation	T	Test result of H0 (OR H1)
Freedom in executing the ideas	2.63	1.063	43.924	Proper
Activities on information categorization	3.26	1.066	54.280	Proper
Knowledge organizing	3.06	0.891	61.224	Proper
Knowledge documenting	2.79	1.022	48.622	Proper
Spent budget to stock knowledge	2.73	1.053	46.203	Proper
Doing your job by colleagues	3.20	1.020	55.860	Proper
Valuing experience and knowledge organization	2.64	1.137	41.385	Proper
Individuals' knowledge assessment	2.71	1.126	42.885	Proper

According to table 7, the highest average relates to item 2 (3.26) while the lowest one is item 1 (2.63). To this end, null hypothesis on knowledge retain is refused. It shows that petrochemical companies are in good status of knowledge sharing. Thus, Hypotheses 6 is supported.

Assessing knowledge management aspects

As seen in research findings section, all research hypotheses are supported and all knowledge management six aspects are in proper status. To answer the second question on the fact in which knowledge aspects, petrochemical companies at South Pars Energy Economic Special Zone are in proper situation, Freedman test is used.

Table 8: Freedman test results to study knowledge management aspects

Aspect	Average rank	Rank
Knowledge recognition	3.25	1
Knowledge sharing	3.16	2
Knowledge utilization	2.93	3
Knowledge retain	2.87	4
Knowledge development	2.65	5
Knowledge learning	2.57	6

Freedman test results indicate that the highest rank belongs to knowledge recognition while the lowest rank is knowledge learning.

Discussion and conclusion

Concerning the importance of knowledge management discussion in today organizations especially oil, gas and petrochemical companies, knowledge management components are measured by Proust, Rob Wermerhardt model in petrochemical companies at South Pars Energy Economic Special Zone as the most important hub of the Middle East economy. Research findings indicate that respondents have assessed the status of knowledge management components as proper in petrochemical companies. A similar research was conducted in Law

Enforcement University and the author concluded that no knowledge management components are suitable. In fact, petrochemical companies had scores higher than average (3) and the most suitable situation in two knowledge recognition and knowledge sharing components. In the meantime, Freedman test results indicated that knowledge management components do not enjoy similar situation. Knowledge recognition is the most suitable one and rank (3.25) followed by knowledge (3.16), knowledge utilization (2.93), knowledge retain (2.87), knowledge development (2.65) and knowledge learning (2.57). this proper situation can help studied companies in technological development, scientific development to improve production quality and quantity, waste reduction, HR attraction, documenting, credits, coding and knowledge distribution and convert them into successful enterprises.

References

Aminpoor, F, (2006). *Knowledge management in organization*. Change Management Conference, Esfahan.

Bakhtiari, H, (2009). *the necessity and importance of knowledge management in information age*. the 1st Executive Management Conference,Tehran.

Becerra-Fernandez, I. (2000) "The role of artificial intelligence technologies in the implementation of people-finder Knowledge management systems", In *AAAI Workshop on Bringing Knowledge to Business*.

Davenport,TH., De long, DW.,& Beers, Mc (1998)."Successful Knowledge Management Projects", *Sloan Management Review*,Vol.10(5),pp.43-57.

Gloet,M. & M. Terziovski. (2004)."Exploring the relationship between km practices and innovation performance", *Journal of Manufacturing Technology Management*, 15(5): 402–409.

Keyvani, Maysam.(2011). *Providing a customer's knowledge management framework in Shargh Cement Company*. M. A. thesis. Tarbyyat Modares University.

Massa, S. & Testa, S. (2009). A knowledge management approach to organizational competitive advantage evidence from the food sector. *European Management Journal,* 27, 128– 141.

Miguel, A, "*Linking properties of knowledge and knowledge network topology with performance*", A dissertation of doctor of philosophy, University of Nebraska, 2007.

Nasr Esfahani, Hassan; Taheri, Mohammad Ali & Goly, Mohammad Reza (2008), studying the essentials and challenges of establishing knowledge management in Iranian National Gas Company, sixth *International Conference on Management*, Tehran

Roding, Allen (1998), *knowledge management in information – based global economy*, translated by Mohammad Hussein Latifi (2007), Tehran, SAMT Publications.

Walczak, S. (2005). Organizational Knowledge Management Structure. *The Learning Organization*,12(4),330-339.

Wiig, k., &Prousak.J.(2009). Knowledge management foundation Arington: TX: Schama Hibbard, J.Knowling what we know. Information weec,october,No.20 P:16 -25.

Study of Barriers to Scientific Collaboration of female Scientifics (Case Study of Iranian Women members of University of Tehran)

Fatemeh Nooshinfar*, Aref Riahi** and Elham Ahmadi***

*Assistant Professor, Knowledge & Information Science, Azad Islamic University,
Branch of Science and Research of Tehran, Iran

**PhD Student Knowledge & Information Science, Azad Islamic University,
Branch of Science and Research of Tehran, Iran
ariahi@ut.ac.ir

***MA Student Knowledge & Information Science of Shahid Chamran University (Corresponding Author),
ahmadielham2012@gmail.com

Abstract

The objective of the present research is to investigate the challenges and barriers of international scientific cooperation from the viewpoint of the female researchers affiliated to Tehran University in co-authored papers and documents in ISI database. The population is 34 consisting of female members of academic board in Tehran University, the persons who have been co-authored on collaborative writings with foreign researchers (and Iranians living abroad). This research findings have shown that political problems and barriers averaged (3/81) are the most effective barriers to international collaboration from the viewpoint of the study population. Cultural variable averaged (2/19) has had the least amount of impact compared with other elements. Besides, among 54 components related to 6 main study variables, 'insufficient mastery over English language or other languages' averaged (4/18) has been the most important barrier to international scientific cooperation or collaboration. 'Religious differences in scientific collaboration with foreign researchers averaged (1/73) has had the least amount of impact. Besides, the results have shown a negative and significant correlation between international collaborative scientific writings of the respondents and efficiency of three political (-0/485), organizational (-0/423) and motivational (- 0/412) variables. In other words, the more the impact of political, organizational and motivational barriers and problems among the respondents responding to the questionnaire, the less the number of international joint authorship.

Keywords: Scientific collaboration, International Collaboration, Challenges and Barriers, Joint Output, Co-authorship, Iranian Women Researchers, ISI database

Introduction

Increasing development of technology in recent decades is the most important axis of development in developed and developing countries; it has a special place in cultural, social, economic and industrial development programs and above all leads to higher living standards and welfare. Undoubtedly, each country's scientific and technical production capacity is the most significant indicator of countries' development. Besides, each country's universities and higher education centers are among the most important centers of science production; they take part in these academic activities to develop the culture and life style of the public life.

Achieving a right understanding of capabilities, facilities, discovering the strengths and weaknesses of the group and team researches and organizing them are very important. In other words, we can state that awareness of barriers to collaborative team researches and resolving them (to promote academic collaborations quantitatively and qualitatively) is very important for different countries' academic policies. This issue is more important for developing countries female researchers (especially Iran).

Here we should mention that discrimination between public and private life and speeding of Modernization process in Modern life have increased the ration of women dedicated to private and family life and limited them more and more to private realms and decreased their independent activities; besides, women until recently have had a minor role in different fields including academic domains.

It is worthy to mention that in the modern era accelerating the pace of the modernization process, creating a distinction between public and private life has increased the presence of women in the private life and limited them to family life and decreased their independency; therefore, women , almost until recently, had little role in various fields including scientific domains.

In Iran where women are deprived of a lot of rights and social affairs under patriarchy, the cultural changes made due to modern elements, the rise of Islamic Revolution in 1357 and the weakness of patriarchy culture have changed it a lot. Extensive participation of women in society especially their presence in education is obvious. As statistics shows, every year more than 50% of the students getting accepted into the universities are women; it gradually makes the presence of women more significant in higher education (MSc and BSc).

Women's high turnout in higher grades of the country's academic community has increased their participation in scientific production at national level, though Iranian women's scientific production growth level has not been at the same level as their presence in academic communities.

Furthermore, we should state the low level of participation at international level and a few collaborative researches among female Iranian researchers and scholars and their foreign counterparts. Muslim female Iranian scholars have often encountered some problems_ at international researches domain _ its rate and intensity have been higher than male researchers ones; academic politicians have paid less attention to above mentioned barriers and problems.

Therefore, the present study intends to investigate the barriers and challenges to female Iranian researchers' academic collaborations in co-compilation with foreign colleagues and the attempt to resolve and remove them in order to cause the development of countries in different academic fields and joining female Muslim Iranian community to academic and research networks at international level.

Methodology and data collecting

This is a practical survey. Women members of academic board are this study's sampling population. They teach at Tehran University and co-compiled academic productions with their foreign colleagues at ISI database. Indexed ISI Database (address http://apps.isiknowledge.com) through the utilization of search methods and formula has shown that 69 academic certificates have been produced by these individuals. Now, 44 female members of academic board of Tehran University have taken part in the above mentioned organization to produce these evidences. Questionnaires have been used to collect data having been sent to 44 ones and all of them have responded this questionnaire. This questionnaire has had 70 closed and 21 open questions. 5-point Likert scale range has been used for each closed question. SPSS has been used to analyze data.

Objective of Research

The main Objective of the present study is Study of Barriers to Scientific Collaboration of female Scientifics (Case Study of Iranian Women members of University of Tehran)

Literature Research

Researchers have done a lot of researches on academic collaboration and the elements affecting them in Iran and other countries. Academic collaboration studies have been dynamic and growing and the number of researches has been increasing every day. Researchers have paid more attention to studying this domain due to the increase made to havehad better quality in most of joint and collaborative researches, besides, researchers' increasing tendency to conducting researches in groups. Below is the review of some studies about this subject.

Biver and Rosen (1978) have proved the large effect of gender, common nationality, motivations and strategies, common culture, religion, language and geographical location on scientific cooperation. Ketza and Martin (1997) have investigated the limitations that may happen doing joint and collaborative research from three financial, time and managerial dimensions. Child and Fakner (1998) have shown reasons such as "Resource Dependence," "learning, reducing risk," "fast market access", "cut costs" and "poor performance of individual activities of against group activities" among the most important reasons for collaboration between organizations and individuals. Matsich and Moonesi (2002) have investigated the factors leading to success in collaboration and classified them in 6 main categories. These factors are: environment-related factors, resource-related factors, and goal-related factors.

Osareh and Wilson (2002) studied cooperation in producing scientific works for Iranian researchers in ISI database during 1995-1999 and compared the results with previous periods. Results showed that scientific articles in 1995-1999 was increase 2.8 fold than previous period and 1990-1994 has two fold growth relative to previous period 1985-1989. End of Iraq-Iran war, improve in economic conditions, change in scientific attitudes and encourage researchers are the main factors for developing scientific cooperation in Iran.

Hara & et.al (2003) identified effective factors for cooperation in their study "scientific cooperation: researchers' view about cooperation and effective factors". They classified these factors in 4 groups: individual adjustment, work relation, motivation and social-technological infrastructures. Royle et.al (2007) stated in their research that geographical proximity and political ties were very influential in shaping scientific cooperation between Chinese researchers.

Kuhen and Robert (2008) have emphasized the rapid scientific development in China with researchers' return from other countries specially North America and Western Europe and its most important is these researchers' scientific communication with other researchers. This research results have shown the positive impact of experience of the researchers living outside china on Chinese researchers' scientific production. Paolin et al (2008) have shown that scientific production of the researchers living in England, France, and Germany has been higher than EU countries and the most important reason is the interest in cooperation with their colonized countries during the years not far away; besides, there has been a lot of similarities between them in terms of language and culture. Nilos (2009) has shown that researchers in Cameroon Africa have had a lot of scientific cooperation with European researchers (especially French researchers and due to language similarities) and one of the most important reasons for scientific cooperation is the need to think and use the developed countries' knowledge and expertise, academic affiliation to them, and the inability to do the research by themselves.

We can propose the following model as a pattern for the barriers related to Iranian researchers' scientific cooperation (especially members of academic board) considering the assumptions and theories regarding scientific cooperation.

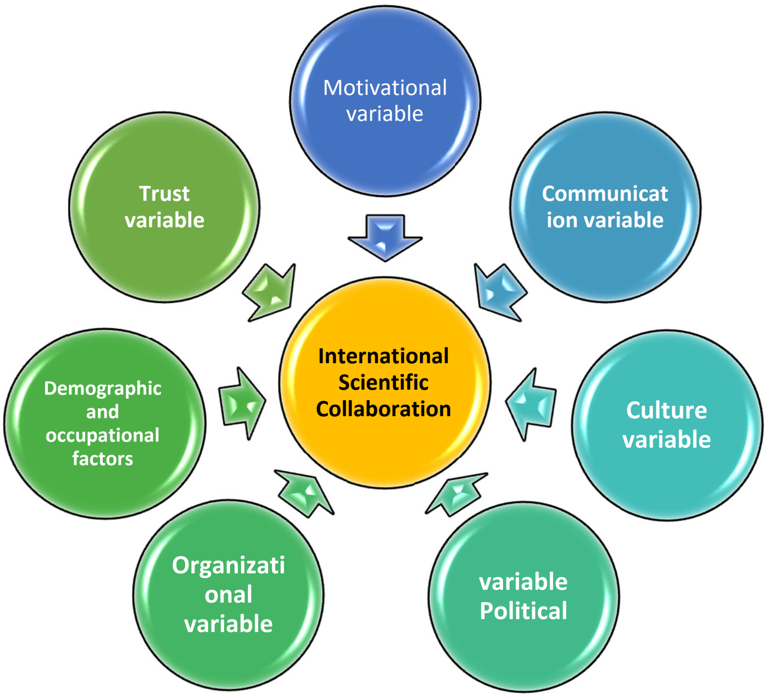


Figure 1. Pattern of barriers and challenges to researchers' scientific cooperation (especially members of academic board)

Operational definitions

Demographic and occupational factors: the present research has pointed to factors such as age, different academic ranks, different statuses of recruitment, different locations of getting degree, and different educational departments.

Motivational variable (individualism): In the present research, unscientific or scientific motivational factors are that category of those factors and material benefits and employment (rank) affecting scientific cooperation. In the present research, unscientific motivational factors with coefficients including employment promotion, financial interests (economic) and personal ownership of pertinent ideas and knowledge. These elements are opposite scientific motivational (social-oriented) elements.

Trust variable: in the present research, trust is the amount of trust of the researchers towards each other; on this basis that others behave as expected and what they say is reliable. Here this variable indicates the researchers' mutual trust and confidence in academic cooperation and their scientific secrecy.

Communication variable: the present research measures the communication variable through two coefficients (1) researchers' communicative ability and (2) communication.

Culture variable: here the culture variable means cultural, religious and lingual differences between different countries researchers, international research culture and scientific cooperation culture with foreign researchers in a scientific society having known as an effective element in international scientific cooperation.

Political variable: in the present research, the political variable is political relations, exchanges and transactions, communications and reactions and in general international relations between countries having known as an effective barrier to international scientific collaborations.

Organizational variable: organizational variable refers to culture, structure and organizational resources (inhibiting factors) including international protocols and agreements, policies and organization's written policies as well as managers' decisions regarding international scientific cooperation.

The research done on Demographic and occupational factors influencing respondents to the questionnaire has shown there has been no significant difference between the researchers' responses from different groups, ages, employment status, the universities from which they have taken the degree, and different academic ranks.

Regarding motivational variable assessed through six items, we should mention that "Dispute between the foreign partners over the order of the name brought in a joint work" with average of 3/29 has been known as the most effective factor. Regarding the variable concerning confidence assessed through 8 items, we should mention that "foreign researchers' insufficient participation in joint research projects" has been known as the most effective factor with the highest average of (3/77). Regarding barriers to communication, it's necessary to mention that "inadequate mastery of English or other languages" has been known as the most important barrier to academic international collaboration with the highest average of (4/11). Regarding cultural and political barriers, we should note that "low academic collaboration culture and teamwork with foreign researchers" with the average of 2/77 and "government intervention and taking improper decisions regarding academic collaboration" with average of 3/84 has been identified as the most effective barrier. These two variables respectively 13 and 6 items have been assessed. Regarding organizational variable _ which has been assessed like cultural variable with 13 items _ it is necessary to mention that "university and organization regulations concerning fellowships, presence in national and international conferences" with average of 3/46 has been considered the most effective organizational factor.

Table1 shows the most important barriers to international scientific cooperation (3 cases) of respondents responding to questionnaire in terms of different variables (together with mean and standard deviation).

Table 1. Barriers to international scientific cooperation

Factors	*The most important barriers to international scientific cooperation*	*Average*	*Standard*
	Dispute over the order of the names in a joint work with the foreign colleague	3.39	1.120
Motivation al variable	Getting a higher grade in individual compilation	2.93	1.147
	Doubt about recording thoughts and opinions in cooperation with foreign colleague	1.485	2.87
	Insufficient cooperation of foreign researchers in doing joint research	3.54	1.137
Trust variable	Work in new and unknown environments	3.41	0.986
	Not having confidence on foreign researchers regarding sharing new ideas	3.09	1.316
	Insufficient mastery over English language and other languages	4.11	1.015
Communic ation variable	Not having access to proper communicative equipments	3.37	1.303

	Inability in having communicative relationship with foreigners	3.10	1.118
	The low level of academic culture of collaboration and teamwork with foreign researchers	2.77	1.377
Culture variable	Doing different parts of research in different countries	2.47	1.208
	Differences of opinion with respect to intellectual property of joint work with foreign partners	2.29	1.129
	Government interventions and making improper decisions regarding scientific cooperation	3.48	1.032
Political variable	International isolations and sanctions	3.66	1.407
	The problems encountered during visa issue for parties to international scientific cooperation	3.52	1.357
	University and organization regulations regarding study opportunities and presence in international conferences	3.46	1.205
Organizational variable	Lack of financial support of the organization's joint researches	3.29	1.342
	The presence of academic bureaucracy	3.11	1.414

Discussion and Conclusions

Researchers have to cooperate with each other considering the present condition of academic community. They can achieve knowledge, skill, sources and facilities through academic cooperation; it is difficult for researchers to achieve them individually. Researchers' familiarity with these benefits and their acceptance of those benefits arising from collective activities will have solved most of the problems and barriers to scientific production.

The results have shown that dispute with foreign colleagues over the order of the names has had the highest influence in scientific cooperation; this can be caused due to unfamiliarity with cooperation culture in doing research in groups. If researchers pay attention to cooperation culture, we won't encounter such problems in doing scientific researches in groups. Therefore we can encourage them more and more to cooperate with each other through eliminating the barriers to scientific cooperation and laying foundation for successful and effective cooperation among scientific community including women in country's universities.

We can also point to insufficient cooperation of foreign researchers in doing joint researches as the most important barriers to trust variable; perhaps this lack of confidence and insufficient cooperation are due to the researcher's costly expenses at international level or due to the terror arising from not having sufficient knowledge versus scientific cooperation internationally.

Insufficient mastery over English language and learning English have been known as the most important barrier to communication. Sufficient mastery over the official language of the world and other foreign languages has been an important factor in communicating with researchers internationally. And as you know insufficient mastery over foreign countries has been known an important barrier to international cooperation. Using English language as the language of education and research not only promotes scientific cooperation internationally but also reinforces scientific work quality. Because studying writers' works from other countries provides access to modern information. Here we propose English language proficiency workshops. The result is the same as the research done by Biwer and Roozn (1997), they know common language effective in scientific cooperation.

The low level of scientific cooperation culture and group work with foreign researchers is the most important barrier to culture. Scientific cooperation often represents colleague researchers' work quality and research groups. In the researches done in groups, cooperation causes them to use each other's mastery, talent and experience. This causes an increase not only in the work quality, but also in the experience and learning the unknown. It should be planned to make cooperation and group work culture between our researchers and foreign ones especially women and they should be encouraged to cooperate scientifically with foreign researchers.

Governmental intervention and making improper decisions regarding scientific cooperation are among the most important political barriers. Special political barriers lead to Iran's scientific backwardness or retardation in a long term including not being in connection with developed countries anymore and made some problems in gaining access to modern knowledge and cooperation with countries' researchers to have access to useful database.

Universities' and organization's regulations regarding study opportunities and presence in international conferences are among the most important barriers to the organizations. Some parts of this problem refer to unfavorable political relations at international level and some parts to administrative affairs and their costs. Lack of communication or weak social, political, economic etc. communication weakness affects scientific cooperation. Iranian scientists are required to have scientific communication to observe developed countries' developments and country's universities do not have a proper structure to establish international cooperation. International cooperation requires financial facilities, unfortunately our country and universities do not consider that.

In general, the results of this study can solve the barriers to academic collaboration at international level. Undoubtedly, individuals can take benefits from academic collaboration. Therefore, it is necessary to take measures not to fall behind advanced communities in order to plan a cooperation culture between researchers especially our country's female scholars to take steps, even a few, towards our country's academic development through utilizing teamwork benefits. The present researcher proposes the following suggestions:

1. Considering academic collaborations especially cooperation with researchers at international level, it is necessary to plan more realistic programs concerning academic communications between different countries, higher educational institutes and research centers at international level.
2. A proper condition for doing academic activities to be provided in groups through supporting researches done in groups at international level to facilitate academic exchange and cooperation between researchers at international level.
3. Difficulty of sending Iranian researchers abroad to continue education or have study opportunities (sabbatical) as well as the problems arisen due to issue visa for both cooperative parties at international level are among the most important political, scientific- research barriers. We propose the universities plan in a way that researchers encounter less problems in study opportunities and be able to keep their scientific cooperation with foreign researchers or continue education abroad. Universities on behalf of Ministry of Science, Research and Technology should contribute more to facilitate sending Iranian researchers abroad to continue education and study opportunities (sabbatical).
4. Universities' administrators should plan properly to develop, target and direct academic collaborations at national and international levels. It can be implemented through raising

funds for research, providing facilities especially for costly researches, creating a proper situation for increasing facilities concerning communication with researchers, allocating special funds for joint research projects, holding conferences and academic ones at national and international levels, etc.

5. Authorities' valuing collaborative academic activities and allocating adequate funds and facilities for these types of activities.

References

Beaver, D. & Rosen, R. (1978). Studies in scientific collaboration – Part I. Professional origins of scientific co-authorship. *Scientometrics*, 1(1), 65-84.

Child, J., & Faulkner, D. (1998). *There are no sources in the current document. Strategies of cooperation*. New York: Oxford University Press, Vol.1.

Hara, N, Solomon, Kim & Sonnenwald (2003). An emerging view of scientific collaboration: scientist' perspective on collaboration and factors that impact collaboration. *Journal of the American Society for Information Science and Technology*, 54 (10), 952-965.

Katz, J.S., & Martin, B.R. (1997). What is research collaboration?. *Research Policy*, 26 (1), 1-18.

Koen, J., & Robert, T. (2008). Chinese researchers returning home: Impacts of international mobility on research collaboration and scientific productivity. *Scientometrics*, 77(2), 309-333.

Matsych, Paul W. , Moonesi, Barbara R.(2002). *Success Factors Collaboration*, translated by Maryam Nazari, Tehran: Iranian Information and Documentation Center; 2002.[Book in Persian]

Nelius, B. (2009). Neo-colonialism and research collaboration in Central Africa. *Scientometrics*, 81(2), 413-434.

Osareh, F. & Wilson, C.S. (2002). Collaboration in Iranian scientific publications, *Libri*, 52(2), 88-98.

Pauline, M. & et al. (2008). Intra-EU vs. extra-EU scientific co-publication patents in EU. *Scientometrics*, 75(3), 555-574.

Indian Journal of Physics: A scientometric analysis

Gayatri Paul and Swapan Deoghuria

Indian Association for the Cultivation of Science, Kolkata, India
libgp@iacs.res.in, ccsd@iacs.res.in

Abstract

Peer reviewed journals in the field of Science, Technology and Medical (STM) are the main vehicles through which scientists publish their research output, communicate opinions and exchange observations. Governments and industries of many countries are investing more money than before in science research. As a result, research output in terms of publications of original research articles has been increased substantially in the recent past. So there is a steady growth in STM journal publishing industries too. But unfortunately journals published by universities and learned societies of third world countries are facing a stiff competition from large commercial publishing houses and those are on the verge of extinction due to merger and acquisitions by large commercial publishers. India is one of the fastest growing nations in terms of research output in science. Over the last few years India is trying to establish herself as a global leader in science. The number of original research articles published by Indian scientists has been increased substantially in the recent past. But very few journals in science and specifically in the subject physics with high impact factor are published from India. As a result Indian scientists have no choice but to use journals with high impact factor published from outside India to publish their research output. Amongst the core physics journals published from India, Indian Journal of Physics (IJP) is one bright exception. It not only survived the stiff challenges from commercial publishers but also excelled in many ways that is clearly visible with steady increase in its impact factor over the last few years and for the year 2012 the impact factor of IJP is 1.785 that is highest for any physics journal published from India and comparable with other well known physics journals published from USA and European countries. Bibliometric and scientometric studies were carried out for individual journals in the past for different purposes. IJP is now in the centre of attention to physicists all over the world because of its reasonably good impact factor and the journal is getting more number of original articles from all over the world. The objective of this study is to throw light on the factors those play the vital role for the improvement of its quality by analysing different bibliometric and scientometric data for IJP. The result of this study may be useful and the measures taken by IJP can be extrapolated to other similar journals published by universities and societies from third world countries for improvement.

Introduction

Periodicals are primary source of information and an important media for communication. They play a major role for communicating the latest research findings through publishing articles containing the current development in any field of knowledge. Information is one of the most important resources for a nation that forms the integral base for its economy. Information is growing out in an exponential rate which is often referred to as information explosion. Periodicals publication is also increasing day by day since the first scientific journal started publication in 1665. Periodicals are the indicators of literature growth in any field of knowledge. The advent of Internet technology has led to changes in the way journals operate, including faster review times, electronic submissions and tracking, and online publications. Online access of scientific literature has brought remarkable changes in the way knowledge is shared and disseminates due to its easy availability.

In this study we have considered Indian Journal of Physics to analyse different scientometric data for a period of ten years (2004-2013) because in recent past the journal has showed a remarkable growth both quantitatively and qualitatively. Indian Journal of Physics started its journey in the year 1926 and it is the oldest physics journal published from India. The journal is the brainchild of Sir C V Raman, the Nobel laureate physicist from India and he was the founder and first Editor. Prof. Raman felt the necessity of a physics journal of his own country India in those early days because it was not easy to communicate and publish original research work by Indians and that was also very time consuming. Prof. J C Bose, Prof. S N Bose and other scientists at that time did not get their due credit for their original research. Many of the original research works of Prof. Raman were published in IJP and the second volume of the Journal published his famous article "A New Radiation", reporting the discovery of Raman Effect. Not only Prof. Raman but other doyens of Indian science like K S Krishnan, K Banerjee, S R Palit were contributed in IJP. IJP is a monthly journal in the field of physical sciences that covers almost all branches of physics namely Astrophysics, Atmospheric and Space physics, Atomic & Molecular Physics, Biophysics, Condensed Matter & Materials Physics, General & Interdisciplinary Physics, Nonlinear dynamics & Complex Systems, Nuclear Physics, Optics and Spectroscopy, Particle Physics, Plasma Physics, Relativity & Cosmology, Statistical Physics. Apart from its good user base, the journal is exchanged with many other journals published by learned societies of other countries. The journal is devoted to the publication of original scientific research results in the form of full papers, short notes and Rapid Communications. It also publishes Review Articles from time to time. The Journal emphasizes both fundamental and applied research work in Physics. The journal also publishes Reviews on books under Book-Reviews section. Proceedings of National and International Symposia held in India and Annual Endowment Lectures of IACS are also published from time to time. In addition, Special issues dedicated to distinguished physicists are also brought out. This journal is abstracted / indexed in SCOPUS, INSPEC, Chemical Abstracts Service (CAS), Google Scholar, Academic OneFile, Indian Science Abstracts, INIS Atomindex, INSPIRE, International Bibliography of Book Reviews (IBR), International Bibliography of Periodical Literature (IBZ), OCLC, SCImago, Summon by Serial Solutions.

Indian Association for the Cultivation of Science (IACS), the host institute and publisher of IJP is the oldest research institute (established in 1876) in India. It is devoted to the pursuit of fundamental research on physical sciences. Prof. Raman worked at IACS during 1907 to 1933 making discovery on scattering of light in1928, which bears his name and that brought the Nobel Prize in 1930. The American Chemical Society designated the Raman Effect as an International Historic Chemical Landmark in 1998 and honoured IACS. Apart from Raman almost all leading scientists at that time worked at IACS. Still it is one of the best performing research institutes in terms of research output, international collaboration and accolades. Till 2008 IJP was published, printed and distributed by IACS. In 2009 IACS took a historic decision to sign a co-publishing agreement with Springer, a leading name in journal publishing industry to delegate the right and license to electronically publish and distribute the SpringerLink Edition, and to distribute the International Print Edition outside of India. From 2013 onwards Springer is printing and distributing both national and international version of IJP.

The co-publishing agreement between IACS and Springer is the major turning point and can be considered as a perfect marriage between a reputed institute and a leading industry. In this paper we have found that it is a remarkable turnaround for IJP since 2009 because after that the journal is doing extremely well in all aspects of a STM journal. IACS and IJP have their reputation in scientific community all over the world and Springer has the strength of its marketing strategy, global presence taking advantage of using latest software and technology. IACS still holds the exclusive copyright of IJP and all editorial decisions and processes are being handled by the editorial board and the editorial office of IJP. International Advisory Board, Board of Editors, Honorary Associate Editors and Editorial office take care of the manuscripts submitted by researchers. IJP is now using the state of the art software provided by Springer for manuscript tracking, reviewer selection and reviewer database. The software is efficient, easy to use and helps to expedite the processes between submission of a manuscript and final decision. Since all editorial processes are controlled by the Editorial team, so there is no chance of compromise of the quality of the journal.

Purpose

Scientific publishing is one of the fastest growing sub-sectors of the media industry. STM market is a stable and reliable field for long-term investments. Considering the vast potential of research output from India, this study may throw light on the prospect of publishing more and more journals in other fields of science from India. Our aim is to find opportunities of publication of new journals following the success routes of IJP. We want to study IJP in the limited period of ten years as a case study as IJP is doing very well in the last few years. We know many factors may influence whether a paper is cited much or little, but these cited numbers are best used to obtain an overview of a researcher's output and overall impact (measured as citation counts per article) of journals in knowledge dissemination.

Objective

The present study has been undertaken with the objective of analyzing the following aspects:

1. Publishing trend
2. Authorship pattern
3. Analysis of citations
4. Affiliated institutes of citing authors
5. Countries of collaborating authors and the collaboration
6. Subject analysis
7. Analysis of the pattern of citing journals

Sources of Information

Indian Journal of Physics, Vol.78 (2004) to Vol. 87 (2013) in both hard copy and soft copy (http://www.iacs.res.in/ijp) is the primary sources of information to collect the data. For the information on citation we have consulted two international online databases namely Science Citation Index (SCI) of ISI Web of Knowledge (http://apps.webofscience.com) and the Scopus Database of Elsevier's SciVerse (http://www.scopus.com/home.url). Besides the Annual Report of the IACS and other related publications of IACS are the main sources of information.

Methodology

The bibliographic records for the analysis are limited to the articles of Indian Journal of Physics published during 2004 -2013. Information regarding citation is collected from WoS and the Scopus Database. These are recorded, tabulated and analysed considering the citation year, cited journals, affiliation of the citing authors and subject area of citation.

Results and Discussion

Table 1. Editors and Impact Factor of IJP

Year	Editor(s)	Impact Factor (IF)
2004	Prof. S. P. Sengupta	-
2005	Prof. J. K. Bhattacharjee	0.072
2006	Prof. J. K. Bhattacharjee	0.195
2007	Prof. J. K. Bhattacharjee	0.265
2008	Prof. S. P. Bhattacharyya	0.175
2009	Prof. S. P. Bhattacharyya	0.226
2010	Prof. D. S. Ray and A. Ghosh	0.291
2011	Prof. A. Ghosh	0.381
2012	Prof. A. Ghosh	1.785
2013	Prof. A. Ghosh	-

Table 1 shows the Impact Factor (IF) of IJP during the period 2005 – 2012 (accessed from www.bioxbio.com/if/html/INDIAN -J-PHYS.html and www.bio21.bas.bg/ibf/IF). The IF of the journal 2004 is not available as it is not included in the said year in Web of Science. When the productivity in terms of the number of articles being published in IJP is concerned, the quantity is going up and at the same time ISI Impact Factor is being maintained, which is treated as the measure for the quality of the articles.

Table 2. Ratio of Articles Published and Cited

Year	Articles Published	Articles Cited	Percentage
2004	245	115	46.93
2005	207	81	39.13
2006	140	49	35.00
2007	106	37	34.90
2008	129	34	26.35
2009	161	96	59.62
2010	185	146	78.91
2011	204	123	60.29
2012	170	113	66.47
2013	189	86	45.50

Table 2 depicts the year wise contribution of articles. It is found that the highest numbers of articles (245) is published in the year 2004; while the least number of articles (106) is brought out in the year 2007. From the Table it is clear that papers published in the year 2010 are cited more (146, 78.91%) during 2010 to March 2014 and papers published in the year 2008 got least citation (34, 26.35%) during the consecutive years 2008 to March, 2014. It is evident that from 2010 onwards the journal is more visible to scientific community and as a result cited more.

Table 3. Frequency of Citations

Year	Citation Year											Total	Avg. Citation per year
	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014		
2004	16	38	64	85	60	31	28	26	22	18	8	396	36.00
2005	-	9	40	56	77	44	20	36	23	29	7	341	34.10
2006	-	-	1	26	22	17	25	22	24	22	6	165	18.33
2007	-	-	-	2	10	18	22	16	14	13	4	99	12.38
2008	-	-	-	-	2	25	26	13	15	13	6	100	14.29
2009	-	-	-	-	-	-	39	50	92	87	46	314	62.80
2010	-	-	-	-	-	-	3	65	390	212	33	703	140.00
2011	-	-	-	-	-	-	-	26	267	317	50	660	165.00
2012				-	-	-	-	-	74	348	82	504	168.00
2013	-	-	-	-	-	-	-	-	-	106	117	223	111.50

Table 3 shows a remarkable growth of average citation per article from the year 2010 onwards.

Table 4. Authorship Pattern in Published Papers

Author	Year										Total	%
	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013		
Single	35	24	31	12	16	23	17	31	17	18	224	12.90
Two	76	60	41	28	43	32	50	49	38	43	460	26.49
Three	63	52	34	33	31	35	44	43	50	55	440	25.34
Four	24	37	14	20	19	21	27	28	22	29	241	13.88
Five	28	24	12	05	11	27	18	32	27	31	215	12.38
>Five	19	10	08	08	09	23	29	21	16	13	156	8.99
Total	245	207	140	106	129	161	185	184	170	189	1736	100

Table 4 shows authorship pattern of the papers published during the period 2004 to 2013. Out of 1736 papers, the maximum number of papers 460 (26.49%) have been contributed by two authors. This is followed by three authors with 440 papers (25.34%), four authors with 241 papers (13.88%), five authors with 215 papers (12.38%) and more than five authors with 156 papers (8.99%). Table 4 also shows that out of 1736 papers single author contributed 224 papers (12.90%) while the rest 1512 papers (87.10%) contributed by the joint authors. It is clear from the above analysis that percentage of single authored papers is less than that of joint authored papers. To determine the extent of collaboration in quantitative terms, the formula given by K. Subramanyam is used. The formula is as follows:

$C = N_m/N_m+N_s$ where,

C= Degree of Collaboration

N_m= Number of multi authored contributions

N_s= Number of single authored contributions

In the present study the value of C is: 1512/224+1512=0.87. This brings out clearly the prevalence of team research in Physics field.

Table 5. Organisation wise Citation of Articles

Inst	Article Publishing Year									
	2004 (%)	2005 (%)	2006 (%)	2007 (%)	2008 (%)	2009 (%)	2010 (%)	2011 (%)	2012 (%)	2013 (%)
Colleges	82 (22.77)	59 (14.82)	63 (22.26)	60 (20.48)	62 (18.96)	89 (19.26)	132 (22.53)	85 (11.97)	97 (19.06)	95 (17.30)
Universities	187 (51.94)	207 (52.01)	152 (53.71)	156 (53.23)	172 (52.60)	203 (43.94)	271 (46.24)	348 (49.02)	269 (52.85)	257 (46.81)
Research Organiza-tions	89 (24.72)	131 (32.91)	63 (22.26)	73 (24.92)	91 (27.83)	169 (36.58)	181 (30.89)	276 (38.87)	142 (27.90)	195 (35.51)
Others	2 (0.55)	1 (0.25)	5 (1.76)	4 (1.37)	2 (0.61)	1 (0.22)	2 (0.34)	1 (0.14)	1 (0.19)	2 (0.36)
Total	360	398	283	293	327	462	586	710	509	549

Table 5 shows that authors from universities cite most followed by research institutes and colleges.

Table 6. Distribution of subjects

Subjects	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	Total
Astrophysics, Atmospheric & Space Physics	12	9	9	12	7	6	18	5	11	11	100
Atomic & Molecular Physics	11	4	6	17	3	7	17	15	7	11	98
Biophysics	4	1	2	3	0	0	0	0	0	2	12
Condensed Matter & Materials Physics	88	104	49	39	88	48	54	23	57	64	614
Nuclear Physics	64	8	9	9	1	56	41	114	14	17	333
Optics & Spectroscopy	16	13	20	4	5	4	17	11	9	6	105
General & Inter-disciplinary Physics	29	46	32	14	7	28	23	24	37	30	270
Nonlinear Dynamics & Complex Systems	5	1	1	1	2	1	1	0	7	19	38
Particle Physics	2	2	7	6	12	9	1	2	8	8	57
Plasma Physics	12	18	4	0	2	2	4	10	15	12	79
Relativity & Cosmology	2	1	0	1	0	0	0	0	4	9	17
Statistical Physics	0	0	1	0	2	0	9	0	1	0	13
Total	245	207	140	106	129	161	185	204	170	189	1736

The subjects of the published articles are categorised as mentioned in the Indian Journal of Physics. Table 6 shows that major contributions are in the field of condensed matter & materials physics followed by nuclear physics and general & interdisciplinary physics. The least contribution is in the field of biophysics as there is no contribution in this field during 2008 to 2012 followed by Statistical Physics. There are specific journals in the field of Biophysics and as a result IJP gets less number of papers in this field. Data in this Table also indicates which branches of physics are attracting more scientists.

Table 7. Geographical Distribution of Contributors

Continents	Year										Total
	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	
Africa	0	1	0	0	0	0	8	8	12	14	43
Asia	358	393	281	291	321	442	430	501	463	444	3924
Australia	0	0	0	0	0	0	0	0	2	3	5
Europe	1	3	1	1	4	13	120	154	19	67	383
North America	1	1	1	1	2	7	28	47	13	21	122
Total	360	398	283	293	327	462	586	710	509	549	4477

From Table 7 we find that IJP gets articles from almost all countries and that definitely establish its status of a true international journal. We notice that from 2010 onwards the contributions from outside India have increased remarkably. Major contributing countries are China (153), USA (102), Germany (79), Iran (74), Egypt (61), Italy (54), Russia (52), Turkey (46) etc.

Table 8. Distribution of Cited Journals

Origin	Year										Total
	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	
National	51	32	11	17	35	109	529	501	342	155	1782
International	167	165	72	63	52	128	86	67	41	21	862
Total	218	197	83	80	87	237	615	568	383	176	2644

Journals in which articles of IJP are cited are categorised as national and international journals. Table 8 provides the number of citation made in national and international journals. It shows that Journals published from India cite more than journals published from abroad during 2004 to 2013.

Conclusions

In this study we have found that almost all physics journals (total 163) cite articles published in IJP. Notable among them are Physical Review (A, B, C, D), Advances in High Energy Physics, Applied Optics, Astronomy and Astrophysics, Canadian Journal of Physics, Chinese Journal of Physics, Euro Physics Letters, Journal of American Chemical Society, Journal of Chemical Physics, Nanomaterials Nanotechnology, Physics Letters, Thin Solid Film, Journal of Physical Chemistry, Indian Journal of Pure and Applied Physics etc. This definitely establishes that IJP is well accepted amongst the physicists from all over the world. It also proves that IJP disseminates quality knowledge as far as physics research is concerned. One of the major advantages of IJP is that it covers almost all frontier areas of physics research compared to other physics journals those focus only on a particular area. Another plus point of getting more contributions from countries where English is not native language of their own is that IJP is not so rigid as far as language and grammar is concerned without compromising the quality of research.

Although this study is limited to IJP but the results may be useful to similar journals published by universities and learned societies of third world countries. We have found that impact factor, contribution from countries other than India, citation of IJP articles in all major physics journals, national and international collaboration have been improved considerably since 2009. Taking advantage of a commercial publishing house like Springer for online publication and wide circulation through a co-publishing agreement, it has now transformed from an obscure science journal to a well known international physics journal. Authors, editors and reviewers of the journal are taking advantage of the use of fully web-enabled online manuscript submission and review systems of Springer.

We consider industry-institute collaboration that started in 2009 between IACS and Springer is definitely a break through for IJP that improves the quality of the journal in dissemination of quality research in physics. Other factors those play vital roles in improving the quality are the reputation of the journal and the publisher, efficient editorial work, use of online manuscript tracking system etc. Global presence of Springer and aggressive marketing help the journal to reach many more scientists.

IJP has taken different positive measures to keep this trend of improved quality of the journal. Archiving of back volumes of the journal (1926-2008) and keeping them on-line have been started by keeping the articles in the Institutional Repository (arxiv.iacs.res.in) of IACS that is OAI compliant and interoperable. Articles that are accepted for publication but not yet assigned an issue and volume number is immediately accessible to researchers through “Online First Articles” section of Springer (link.springer.com/journal/12648/onlineFirst). This has increased the chance of getting more citations of their works. Editorial office has extended support to the authors to improve the language of the articles that originate from the countries where English is not native language of their own.

References

Arkhipov, D. B. (1999). Scientometric analysis of Nature, the journal. Scientometrics, 46(1), 51-72.

Das, A. and Sen, B. K. (2001). Indian journal of medical research: an analysis of citation pattern. *ILA Bulletin*, 37, 9-12.

Garg, K. C. and Kumar, S. (2010). An analysis of the citation pattern of Indian Science journals indexed by SCIE. *Annals of Library and Information Studies*, 57, 365 372.

Kumar, A. and others (2008), Pramana – Journal of Physics: A scientometrics analysis. *Annals of Library and Information Studies*, 55, 201-211.

Mahapatra, G. (1994). Indian library and information journals: a bibliometric analysis of the rate of citation and their characterises, *Library Review*, 43, 41-46.

Paul, G. and Das, A. (2012). Impact of Indian Journal of Physics in knowledge dissemination. In *Information and Knowledge Dissemination: Present Status and Future Direction* (pp. 173-181). Kolkata: Allied Publishers.

Ramesh L S R C V and Nagaraju, A V S S (2000). Citation analysis of the Indian Journal of Information, Library and Society. Indian Journal of Information, Library and Society, 13 (3-4), 171-179.

Singh, J. K. (2014). A scientometric analysis of “Indian Journal of Pure and Applied Physics” (2006-2010): a study based on Web of Science. *Research Journal of Library Science*, 2(1), 7-12.

Sivasekaran, K and Ragavan S. S. (2014). Journal of Astrophysics and Astronomy : A Bibliometric Study. e-Library Science Research Journal, 2(6), 1-5.

Subramanian, K. (1983). Bibliometric studies of research collaboration: A review. Journal of Information Science, 6(1), 33-38.

Conflicts of Interest Statements on Biomedical Papers

Grant Lewison and Richard Sullivan

King's College London, Research Oncology, Guy's Hospital, Great Maze Pond, London SE1 6RT, UK
grantlewison@aol.co.uk, richard.sullivan@kcl.ac.uk

Abstract

This paper examines the varying prevalence of conflict of interest (COI), and "no conflict" (NC), statements on biomedical research papers, which are increasingly being required by journal editors. They are important as they may detract from the perceived objectivity of the results if the authors are in the pay of commercial companies. Numbers and percentages of these statements in the Web of Science (WoS) increased from 2009 to 2012 but both started to decline in 2013 and are still only a few percent of the total. They occur most frequently on papers by north Americans and western Europeans, but rarely for authors from east Asia. One consequence of the appearance of COI statements is that the WoS mistakenly includes companies who have given money to some of the researchers for unrelated work among the sponsors listed among the Funding Organizations (FO), and this will distort the analysis of the funding of the research being reported in some of the papers and appears nearly to double companies' apparent tally of papers. However, it appears that many COI statements are excluded from the WoS because they are printed separately from the acknowledgement section of the paper.

Introduction

There is now quite an extensive literature on the problems that can arise when authors of research papers have a financial involvement with companies who may have a commercial interest in the results described. Since 1990, the numbers of papers in the Web of Science (WoS) with *conflict*-of-interest* (COI) in their title and that concern biomedical activity has increased dramatically, see Figure 1. Much of this literature deplores the situation that has arisen, where links between pharmaceutical (and sometimes medical device) companies and supposedly objective researchers have become pervasive so that there is bias in the literature on clinical trials and the public trust in science is eroded (Kirkpatrick *et al.*, 2012; Steinbrook & Lo, 2012; Bariani *et al.*, 2013; Gasparyan *et al.,* 2013; Vasconcelos *et al.*, 2013). [There is also a large literature on other aspects of conflict of interest, notably in the financial system, where advice on investments can be tainted by hidden assets.] However, conflicts of interest not only affect researchers and their papers, but journal editors (Smith *et al.,* 2012; Qureshi *et al.*, 2012; Bosch *et al.*, 2013) and publishers who may depend on the lucrative sale of reprints, especially reports of clinical trials sponsored by companies (Lundh *et al*, 2010).

Many of the papers have examined the COI requirements stated by journals in their instructions to authors (Rowan-Legg *et al.*, 2009; Alfonso *et al*, 2012; Khurana *et al.,*, 2012). The conclusion seems to be that most journals require such statements, both of the sponsorship of the research being described (funding sources) and any financial or non-financial ties between the authors and industry. However, examination of actual practice in particular journals or groups of them suggests that this requirement for COI statements is not being adhered to (*v.i.*). There are corresponding problems in the writing of clinical guidelines, and even Cochrane Reviews, where COIs could colour their recommendations for clinical diagnosis and treatment (Kesselheim *et al.*, 2012; Khalil *et al.*, 2012; Langer *et al.*, 2012; Norris *et al.*, 2012). Some papers have examined individual journals in order to determine the prevalence of COI statements on their papers, and to compare this with the journals' stated policy (Blum *et al.*, 2009; Forbes, 2011; Kesselheim *et al.*, 2012; Das *et al.*, 2013).

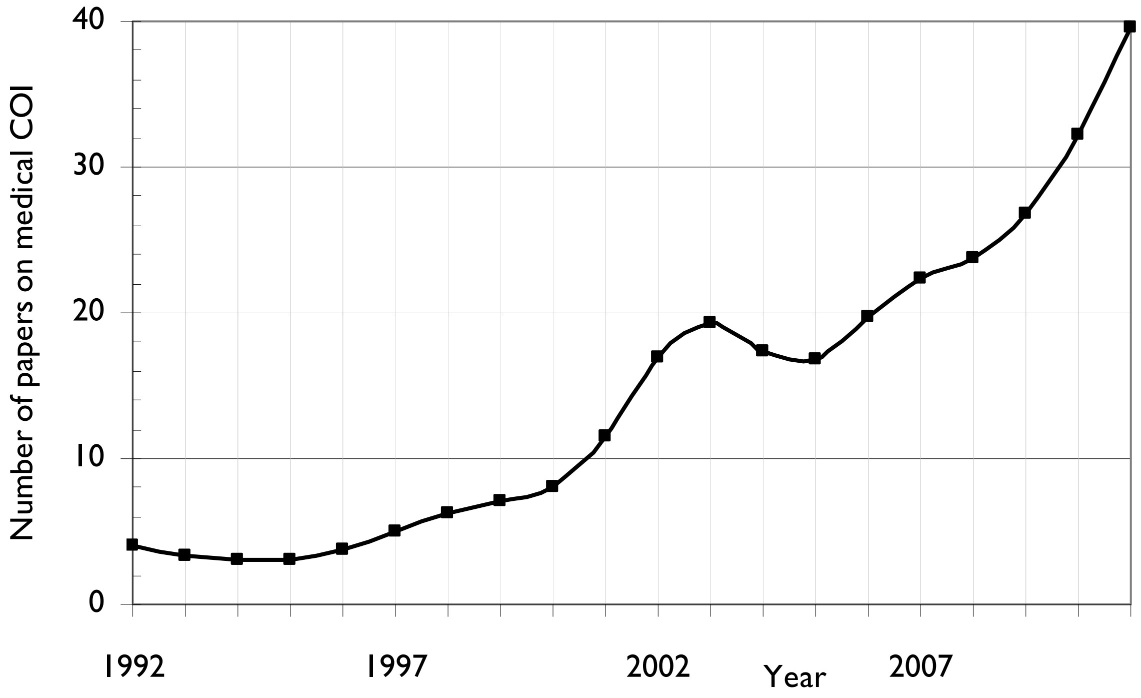


Figure 1. Rise of the numbers of papers in the Web of Science on medical conflict of interest, five-year running means

Declarations of potential conflicts of interest can take several forms. The most common are when authors have undertaken consultancy work, or have spoken on behalf of a company and/or received honoraria or fees for some other activity such as serving on an advisory board. Some authors declare that they hold stock (or shares) in a company, hold patents or receive royalties. However negative statements of “no conflict” (NC) may occur, and sometimes these appear alongside COI statements for some of the authors of a paper.

Since late 2008, the Web of Science (WoS) has routinely included details of financial acknowledgements (and personal ones) in two searchable fields, FO (funding organisation) and FT (funding text), where they occur on a paper. We have used information on funding organisations on several occasions to identify the sources of support for a research portfolio (Lewison & Markusova, 2010; Lewison & Roe, 2012); this has become a relatively routine aspect of research evaluation and may show a research group’s success by how often it obtains external support for its work and from which sources (Lewison, 2003; Rigby, 2013). Because of the aim of governments and charities to make the research they support lead to practical benefit, the involvement of commercial companies in the further development of this research is often seen as desirable. So there is an additional reason to determine how much industrial support has been provided to public-domain research. The data in the FO field (which when downloaded to file appears in a column headed FU) can be used for this purpose.

However, we happened to notice that the list of commercial funders sometimes included companies that had been mentioned in a COI statement that was reproduced in the Funding Text (FT) field (which, when downloaded, is headed FX). It appeared that some of the companies credited in this way were not in fact supporters of the research being reported, but were merely listed as having had financial (or other) links with one or more of the paper authors. This could clearly distort the analysis of commercial funding for research, and also could artificially boost the number of research papers that a company could appear to have supported. We therefore began to investigate how often such COI statements appeared on published biomedical papers, primarily to correct the data in the FU column for our analysis of funding sources. It rapidly became apparent, however, that the frequency of COI statements (complemented by NC ones) was of interest in its own right and could shed light on current practice.

We therefore embarked on a large-scale study of the presence of COI (and NC) statements on journals and papers covered by the WoS (Science Citation Index Extended) during the five years, 2009-13, when inclusion of acknowledgements would have been effectively complete. Since our main concern was with commercial influences on biomedical research and the practice of medicine, we limited the study to biomedical papers and examined the influence of various parameters – the nationality of the authors, the characteristics of the journals and the year of publication – on the prevalence of COI and NC statements. We also looked at the numbers of papers acknowledging support from the top 10 pharmaceutical companies (ranked by R&D spend), and by how much these numbers were inflated by the inclusion of papers where the company had had links with one or more authors but had not funded the research.

Methodology

We first identified and isolated the biomedical papers (articles and reviews only) in the WoS for the five years, 2009-13 by means of a special filter based on address words or contractions (Lewison & Paraje, 2004), such as *allerg*, biochem, canc, dermatol, endocrin*, family, Glaxo*, hlth*. These numbered 2,879,698 in total, and an analysis was made (with the standard WoS software) of the journals in which they were published and the countries of their authors.

Next, we took a large sample of papers with pharmaceutical companies listed among the funding sources and parsed the acknowledgement full texts to see which words occurred most frequently that might be indicative of a possible COI statement. These were individually checked to ensure that they were not used to describe funding for the research being described in the paper. The words that remained, and that appeared to indicate that an author had been retained in some capacity by a company or received some form of payment, were as follows:

> ADVISORY-BOARD or (CONSULT* not CONSULTATION*) or FEES or HONORARI* or LECTURE* or PATENT* or PAYMENT* or ROYALTIES or SERVED or SERVES or SERVING or SHAREHOLDER* or SHARES or (SPEAK* not SPEAKS) or STOCK or STOCKHOLDER

This was then used as the filter for COI statements. In parallel, a simple filter was developed for NC statements, as follows:

> NO-CONFLICT* or NO-POTENTIAL-CONFLICT*

and these two filters were applied to the FT field for biomedical papers in the same years. This yielded 65,001 papers with a COI statement, 38,506 with an NC statement, and 91,760 with either one (or both). The journals and the countries of their authors were analysed similarly for these groups. It was immediately apparent that COI or NC statements were very much the exception among biomedical papers – only 3.3% of all biomedical papers had one.

The 10 pharma companies were the ones listed in Table 1. This table gives their country, a code used in the tables and figures that follow (based on our thesaurus of funding bodies), and the names of their subsidiaries whose research spending would be included with that of the parent in the EU Industrial R&D Scoreboard tables (EU, 2013). The numbers of “their” papers were determined both from the presence of their names in the address field (AD, implicit acknowledgements) and in the funding organisations field (FO, explicit acknowledgements). However, their presence in the funding text field (FT) together with one or more of the terms in the COI statement filter (*v.s.*) argued that these papers should be deducted from the total number papers with explicit (FO) acknowledgements to give a reduced total.

Table 1. List of the 10 top pharmaceutical companies (ranked by R&D expenditures) with their codes and subsidiaries.

Company	*ISO*	*Code*	*Subsidiaries*
Roche	CH	HLR	Chugai Genentech Ventana
Novartis	CH	NVT	Alcon Chiron Ciba-Geigy Genoptix Sandoz
Merck (US)	US	MRK	Benyu Frosst Meriel MSD Organon Schering-Plough
Johnson & Johnson	US	J J J	Alza Centocor Cordis Crucell Depuy Ethicon Independence-Technology Janssen Lifescan Noramco Orapharma Ortho-Cilag Penaten Peninsula-Pharma Pricara Scios Tasmanian-Alkaloids Tibotec Transform-Pharma
Pfizer	US	PFZ	Alacer King-Pharma Pharmacia Searle Sugen Upjohn Warner-Lambert Wyeth
Sanofi-Aventis	FR	SLU	Aventis Genzyme Hoechst Marion-Roussel Medley Rhone-Poulenc Sanofi Synthelabo Uclaf Zentiva
GlaxoSmithKline	UK	GSW	
Eli Lilly	US	LLL	Icos
AstraZeneca	UK	ZAT	Ardea-Biosci Arrow-Therapeut Kudos-Pharma Medimmune Spirogen
Abbott Laboratories	US	ABB	Abbvie Advanced-Medical-Optics Facet-Biotech Knoll Solvay-Pharma

Since it is the editorial policies for COI or NC statements of the individual journals that are of primary concern – what they say and how well they are enforced – the main analysis was of the different journals. Altogether, there were just over 10,000 journals with at least one biomedical paper in 2009-13, but many of them had so few that analysis was not worth while. Attention was therefore focussed on the 5800 journals with at least 20 papers over the five year period, and for these journals the following parameters were investigated:

- research level, on a scale from clinical = 1.0 to basic = 4.0 (see Lewison & Paraje, 2004);
- subject area (if a specialist publication);
- country of publication;
- identity of the publisher or publishing group.

The country of publication and name of the publisher are contained in the WoS data and can be downloaded, and then matched to the journal names. However sometimes there is a difference between the formal name (which is downloaded when individual papers are downloaded) and the name listed in the results of analysis of source titles (which omit any commas, hyphens and ampersands in the title). We were able to ascertain the name of the publisher for about 62% of all the 10,068 journals, containing over 96% of the biomedical papers; most of the remaining journals were not biomedical in character but just contained a few biomedical papers. For the 27 leading publishers (whose journals often gave name variants) we assigned individual trigraph codes and determined their performance. The country of publication was not always evident, but we determined it for 95% of the papers.

Results: statement percentage presence (SPP)

There was a steady increase from 2009 to 2012 in the percentage of papers with a COI statement from 1.2% in 2009 to 3.1% in 2012, but then a decrease to 2.8% in 2013. Initially we thought that this might have been because of the greater number of papers from China and South Korea (which, as we show later, tend to have relatively few such statements), but this was not the case as almost all countries showed a decline in both numbers and percentages of statements in 2013 compared with 2012. Reviews, which account for about 9% of all the biomedical papers in the five years, average 3.1% of papers with COI statements compared with 1.8% for research articles, showing that more of their authors have something to declare.

The countries of the authors of the biomedical papers, on an integer count basis, are shown in Table 2 for 16 leading countries, with the numbers and percentages of COI statements. This shows that authors from western European and north American countries have relatively many more of them than do authors from the four Asian countries. This is probably mainly caused by pharma company papers having a greater presence among the biomedical papers from western European and north American countries, see Table 3. However Canada ranks higher and Japan, lower, in Table 2 than would be expected from their positions in Table 3. The correlation with the International Transparency Index, http://www.transparency.org/, is also quite good, suggesting that a more transparent culture exists with respect to declarations of interest on biomedical research papers in western Europe and north America.

Table 2. Biomedical research outputs from 16 leading countries, 2009-13, and numbers and percentages of papers with a COI statement (ranked by this percentage).

Countries	*ISO*	*All BM*	*COI*	*%*		*Countries*	*ISO*	*All BM*	*COI*	*%*
Canada	CA	141578	6211	4.39		Australia	AU	109015	3628	3.33
Switzerland	CH	65597	2847	4.34		Italy	IT	145466	4479	3.08
Netherlands	NL	98482	4156	4.22		Spain	ES	103837	3038	2.93
UK	UK	253175	10126	4.00		Brazil	BR	85542	1112	1.30
USA	US	941887	37502	3.98		Japan	JP	192956	2193	1.14
Sweden	SE	60462	2312	3.82		S. Korea	KR	99896	828	0.83
Germany	DE	227717	8324	3.66		India	IN	85574	664	0.78
France	FR	141888	4757	3.35		China	CN	270864	1757	0.65

Table 3. Comparison of percentages of biomedical papers, 2009-13, with numbers of papers from 80 large pharma companies (80 p) for 12 leading countries, integer counts.

Countries	*ISO*	*All BM*	*80 p*	*%*		*Countries*	*ISO*	*All BM*	*80 p*	*%*
Switzerland	CH	65597	5129	7.82		Canada	CA	141578	3854	2.72
Sweden	SE	60462	3372	5.58		Japan	JP	192956	4986	2.58
UK	UK	253175	11483	4.54		Brazil	BR	85542	776	0.91
USA	US	941887	37562	3.99		India	IN	85574	594	0.69
Germany	DE	227717	8985	3.95		S. Korea	KR	99896	679	0.68
Netherlands	NL	98482	3287	3.34		China	CN	270864	1619	0.60

We turn now to the results for the various journals. First, we examine the effect of research level in Figure 2. There is a big difference between clinical journals and basic ones as one might expect because conflicts of interest are more likely to occur where drugs are being trialled, or patients treated in other ways, than in research that is some way from application. The peak of SPP is in the RL range 1.2 to 1.4, and it is primarily due to five drug journals, all published in New Zealand by ADIS International Ltd, and with SPP above 30%: *Drug Safety, Drugs, Clinical Drug Investigation, CNS Drugs* and *American Journal of Cardiovascular Drugs.* The majority of journals in the clinical investigation to basic research category (RL from 3.0 to 4.0) have a much smaller SPP, and two thirds of them have a value less than 1%.

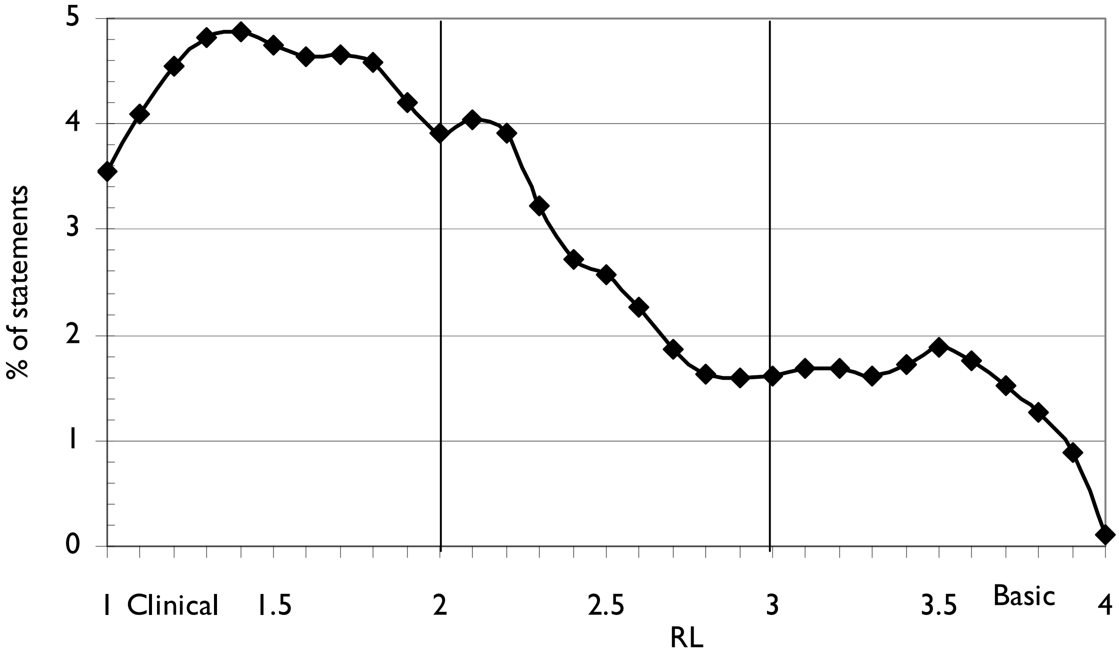


Figure 2. Percentage of COI or NC statements on biomedical papers, 2009-13, and variation with journal Research Level (1.0 = clinical, 4.0 = basic), smoothed curve.

Table 4 shows the SPP (and mean RL value) of some sets of journals: specialist cancer, cardiology, diabetes and surgery journals, ones in the *Lancet* and *Nature* groups (the latter now far more numerous), and ones describing themselves as the *American Journal of…* , the *British Journal of …* and the *International Journal of …* There are some notable differences, with the four *Lancet* journals being the most clinical and having a rather high SPP, and the 43 *Nature* journals being the most basic and having the least SPP. It is perhaps surprising that surgery journals, being very clinical with mean RL = 1.26, have such a low SPP, but perhaps surgeons are not so likely to be in the pay of pharmaceutical companies as are physicians.

Table 5 shows an analysis by publishing country for the leading such countries. The ranking is again similar to that in Tables 2 and 3. Germany (mainly Springer) and the Netherlands (mainly Elsevier) are the leading publication countries in continental Europe.

Table 4. Ranking of different categories of journals by SPP, 2009-13.

Journal type or category	*N (J)*	*Papers*	*Mean RL*	*COI or NC*	*% COI or NC*
LANCET ….	4	2858	1.26	717	25.09
DIABETES	35	16445	1.66	1457	8.86
BRITISH ….	25	23770	1.59	2027	8.53
CARDIO / HEART	122	69437	1.49	4665	6.72
AMERICAN ….	97	73208	1.80	3458	4.72
CANCER	212	128767	1.93	4590	3.56
BMC ….	52	38461	2.18	1123	2.92
INTERNATIONAL ….	408	97102	1.94	2434	2.51
SURGERY	193	132044	1.26	3201	2.42
NATURE ….	43	17189	2.88	207	1.20

Table 5. Ranking of different publication countries with at least 40 journals (Jnls) by SPP, 2009-13.

Country	*ISO*	*Jnls*	*%*		*Country*	*ISO*	*Jnls*	*%*
Canada	CA	55	4.28		France	FR	93	1.54
USA	US	2732	3.83		Germany	DE	332	1.54
United Kingdom	UK	1220	3.76		Poland	PL	78	0.77
Switzerland	CH	154	2.26		India	IN	62	0.76
Italy	IT	76	2.13		Brazil	BR	53	0.60
Netherlands	NL	418	2.02		Turkey	TR	54	0.50
Ireland	IE	47	2.01		S. Korea	KR	47	0.50
Japan	JP	84	1.68		China	CN	42	0.37
Spain	ES	47	1.61					

Results: pharma company papers

We investigated how many papers included one or more of the company names (see Table 1) in the address field (AD, indicative of implicit financial support) or in the funding organization field (FO, indicative of explicit financial support), or both. We also deducted the papers where there was a word showing a potential conflict of interest (see the list in the second paragraph of the methodology section) and one or more of the company names in the funding text (FT). Inspection of a sample of individual papers showed that this almost always meant that the named company had been paying one or more of the authors and that consequently the inclusion of the company in the FO field was incorrect. The results for the ten companies are given in Table 6. A graph of the number of papers for each of the ten (corrected for the presence of COI statements involving the company) against their total R&D expenditure in the quinquennium two years earlier (*i.e.*, from 2007-11) is shown in Figure 3.

Table 6. R&D expenditure by 10 leading pharma companies, 2007-11, € M, and their presence in the WoS: AD = addresses; FO = funding credits, with and without COI statements, and corrected total number of papers in 2009-13.

Company	*Code*	*Spend*	*AD*	*FO*	*Total*	*FOxCOI*	*Total xCOI*
Roche	HLR	31259	6136	12976	17707	5644	10762
Novartis	NVP	27762	5749	17722	21338	7679	11881
Merck (US)	MRK	23365	6385	18963	23452	7609	12667
Johnson & Johnson	J J J	26508	3941	12396	15008	4427	7443
Pfizer	PFZ	30475	7752	24391	29175	11595	17182
Sanofi-Aventis	SLU	22925	2530	12624	14180	5234	7071
GlaxoSmithKline	GSW	21095	5392	14043	17285	6419	10192
Eli Lilly	LLL	15684	2969	10815	12487	4721	6803
AstraZeneca	ZAT	17034	4462	10385	13133	4609	7762
Abbott Labs	ABB	11528	2544	8335	10126	3002	5046

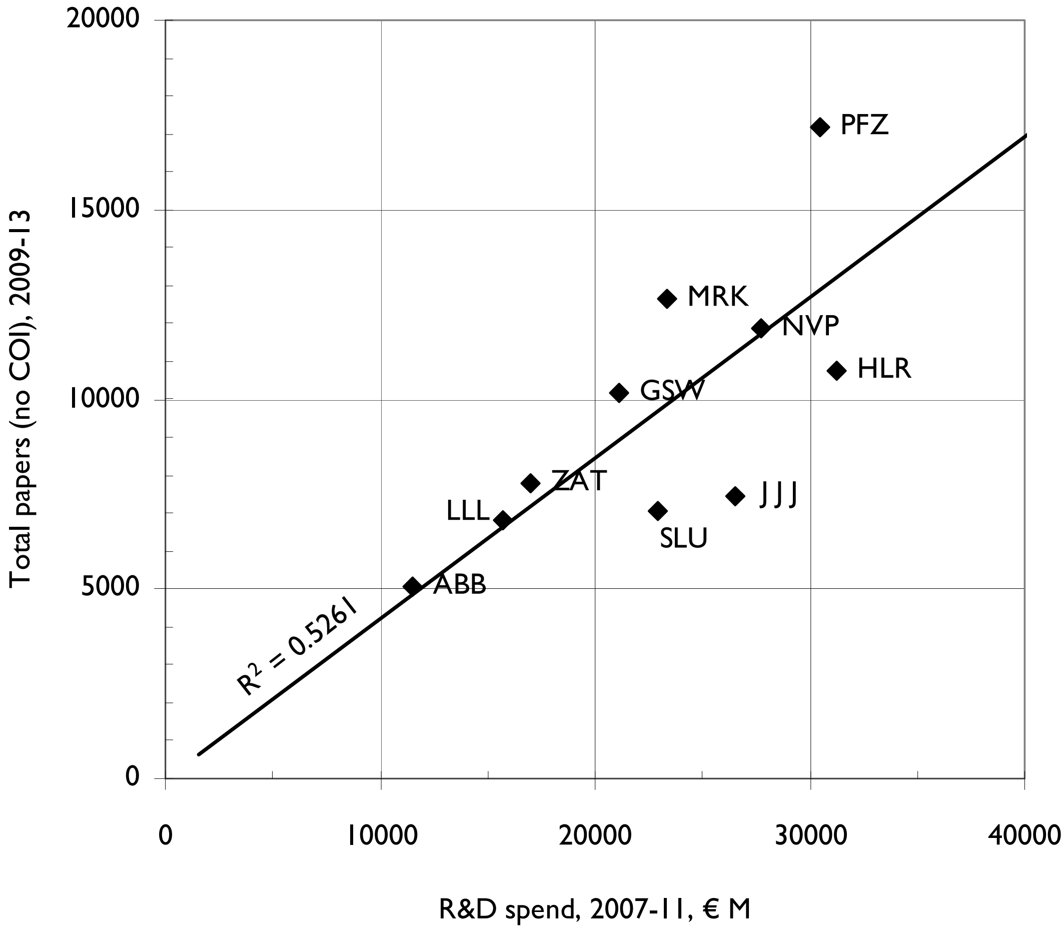


Figure 3. Company research outputs, 2009-13, excluding papers with a COI statement, compared with total R&D expenditure in 2007-11.

The figures in Table 6 show that the reduction in the number of papers acknowledging one of the ten companies because of the presence of a COI statement is large, averaging 42% with standard error of the mean 1.3%. On average, 49% of the company papers have the name in the address field, and 62% include it among the funding organizations, and 12% show the company name in both fields. These results show that it is important to include implicit acknowledgements, and remove papers attributed to a company if a COI statement naming the company is present, when funding analyses are being performed.

Discussion

We wondered if the reduction in SSP observed in 2013 was a statistical fluke, or whether it was the beginning of a trend for authors to include fewer COI statements on their papers, despite the requirements of an increasing number of journals. An analysis of the data for the first four months of 2014 showed that COI statements occurred on 2532 biomedical papers out of a total of 134,886, or 1.9%. This is substantially lower than the figures for 2012 and 2013 which were 3.1% and 2.8%. So it certainly looks as if the SSP is now on a definite downward trend.

We also thought it desirable to check with some journal editors why there appeared to be so few COI statements on the papers in some clinical journals where more might have been expected in view of the clear instructions to authors to provide one. The responses we received (about half of the 14 who were polled) all said that they expected such statements to be provided and expressed surprise that our data showed such a low prevalence. There were a few cases where the lack of a COI had been picked up by a reviewer, and a handful of papers had been rejected or even withdrawn because of this, so evidently it was occasionally of importance – although it appeared that plagiarism was sometimes more of a problem. We examined some recent issues of one journal, the *British Journal of General Practice*, and found that the papers did indeed all have a clear COI or NC statement, but that this was not incorporated in the Funding Text of the WoS record. It thus appears that the SPP may be unduly low, and its declining value in the last two years may be an artefact because COI and NC statements are appearing separately from the formal acknowledgement. The number of corrections to the Funding Organization lists in the WoS may therefore be an under-estimate.

Acknowledgements

We are grateful to Philip Roe of Evaluametrics Ltd who provided the macros that were used in several of the analyses. The authors have no relevant competing interests to declare.

References

Alfonso, F Timmis, A Pinto, FJ Ambrosio, G Ector, H Kulakowski, P et al. (2012) Conflict of interest policies and disclosure requirements among European Society of Cardiology National Cardiovascular Journals *European Heart Journal* 33(5), 587-594

Bariani, GM Ferrari, ACRDC Hoff, PM Krzyzanowska, MK & Riechelmann, RP (2013) Self-Reported Conflicts of Interest of Authors, Trial Sponsorship, and the Interpretation of Editorials and Related Phase III Trials in Oncology *Journal of Clinical Oncology* 31(18), 2289-2295

Blum, JA Freeman, K Dart, RC & Cooper, RJ (2009) Requirements and Definitions in Conflict of Interest Policies of Medical Journals *JAMA-Journal of the American Medical Association* 302(20), 2230-2234

Bosch, X Pericas, JM Hernandez, C & Doti, P (2013) Financial, nonfinancial and editors' conflicts of interest in high-impact biomedical journals *European Journal of Clinical Investigation* 43(7), 660-667

Das, KK Vallabha, T Ray, J & Murthy, PSN (2013) Conflict of Interest - Serious Issue on Publication Ethics for Indian Medical Journals *Journal of Nepal Medical Association* 52(6), 357-360

EU (2013) *The 2013 EU Industrial R&D Investment Scoreboard*. EUR 26221 EN, ISBN 978-92-79-33743-7 (print) Available at: http://iri.jrc.ec.europa.eu/scoreboard13.html Downloaded 10 June 2014.

Forbes, TL (2011) Author disclosure of conflict of interest in vascular surgery journals *Journal of Vascular Surgery* 54), 55S-58S

Gasparyan, AY Ayvazyan, L Akazhanov, NA & Kitas, GD (2013) Conflicts of interest in biomedical publications: considerations for authors, peer reviewers, and editors *Croatian Medical Journal* 54(6), 600-608

Kesselheim, AS Wang, B Studdert, DM & Avorn, J (2012) Conflict of Interest Reporting by Authors Involved in Promotion of Off-Label Drug Use: An Analysis of Journal Disclosures *PLOS Medicine* 9(8)

Khalil, B Aung, K & Mansi, IA (2012) Reporting Potential Conflicts of Interest among Authors of Professional Medical Societies' Guidelines *Southern Medical Journal* 105(8), 411-415

Khurana, G Henderson, S Walter, G & Martin, A (2012) Conflict of Interest and Disclosure Policies in Psychiatry and Medicine: A Comparative Study of Peer-Reviewed Journals *Academic Psychiatry* 36(1), 17-22

Kirkpatrick, JN Kadakia, MB & Vargas, A (2012) Management of Conflicts of Interest in Cardiovascular Medicine *Progress in Cardiovascular Diseases* 55(3), 258-265

Langer, T Conrad, S Fishman, L Gerken, M Schwarz, S Weikert, B et al. (2012) Conflicts of Interest Among Authors of Medical Guidelines *Deutsches Arzteblatt International* 109(48), 836-842

Lewison, G (2003) The publication of cancer research papers in high impact journals *Aslib Proceedings*, 55(5,6), 379-387

Lewison, G & Paraje, G (2004) The classification of biomedical journals by research level *Scientometrics* 60(2), 145-157

Lewison, G & Markusova, V (2010) The evaluation of Russian cancer research *Research Evaluation*, 19(2), 129-144

Lewison, G & Roe, P (2012) The evaluation of Indian cancer research *Scientometrics*, 93(1), 167-181

Lundh, A Barbateskovic, M Hrobjartsson, A & Gotzsche, PC (2010) Conflicts of Interest at Medical Journals: The Influence of Industry-Supported Randomised Trials on Journal Impact Factors and Revenue - Cohort Study *PLOS Medicine* 7(10)

Norris, SL Burda, BU Holmer, HK Ogden, LA Fu, RW Bero, L et al. (2012) Author's specialty and conflicts of interest contribute to conflicting guidelines for screening mammography *Journal of Clinical Epidemiology* 65(7), 725-733

Qureshi, J Sud, A & Vakil, N (2012) Funding source and conflict of interest disclosures by authors and editors in gastroenterology specialty journals revisited *Alimentary Pharmacology & Therapeutics* 35(6), 690-695

Rigby J (2013) Looking for the impact of peer review: does count of funding acknowledgements really predict research impact? *Scientometrics* 94(1), 57-73

Rowan-Legg, A Weijer, C Gao, J & Fernandez, C (2009) A comparison of journal instructions regarding institutional review board approval and conflict-of-interest disclosure between 1995 and 2005 *Journal of Medical Ethic*s 35(1), 74-78

Smith, E Potvin, MJ Williams-Jones, B (2012) Accessibility and transparency of editor conflicts of interest policy instruments in medical journals *Journal of Medical Ethic*s 38(11), 679-684

Steinbrook, R & Lo, B (2012) Medical Journals and Conflicts of Interest *Journal of Law Medicine & Ethics* 40(3), 488-499

Vasconcelos, SMR Cassimiro, MC Martins, MFM & Palacios, M (2013) Addressing conflicts of interest in the research paper: a societal demand in contemporary science? *Brazilian Journal of Medical and Biological Research* 46(12), 1007-1013

Structure and Evolution of Library and Information Science in the top Countries of Middle East in terms of Scientific Productions during the years of 1992-2012

Hamideh Asadi* and Mahsan Poorasadollahi**

*Master Student of Scientometrics, Library and Information Science Department, University of Tehran, Enghelab sq. and Senior Librarian of Periodical Section, Encyclopaedia Islamica Foundation, Keshavarz Boulevard, Felestin Street, Tehran (Iran)
asadi1366@gmail.com

**Scientific Board Member and Senior Librarian of Cataloging Section, Encyclopaedia Islamica Foundation, Keshavarz Boulevard, Felestin Street, Tehran (Iran)
mpoorasadollahi@gmail.com

Abstract

This study decides to represent the structure and evolution of Library and Information Science in the top countries of Middle East in term of scientific productions during the years of 1992-2012. Title words or phrases analysis is a technique used to indentify structure and derived from WordStat software. Co-word is another technique applied to investigate the evolution and VoSviewer is software which used for this purpose. Dendrogram (hierarchical clustering) indicates that Library and Information Science has been constructed of three main branches: Library Science, Information Science and Technology, Bibliometric and Scientometric. Each branch contains several sub branches. The maps (Label View and Density View) provide insight into the evolution of Library and Information Science, visualize the terms used in each decade and thereby reveal that how Library and Information Science has been changed and developed during these periods. Total of these analysis have been carried out based on investigating around the 1000 article extracted from Web of Science database and methodology applied in this research represent a new approach to study scientific disciplines whereby the structure and evolution of a field can be investigated.

Keywords: cognitive structure, evolution, word analysis, co-word analysis, LIS

Introduction

Since World War II, the Scope and volume of scientific research has dramatically increased which has led to the growth of human knowledge that is usually reflected in scientific output of researchers such as journal articles, conference papers and patents (He, 1999). Considering the fact, Price (1963) announced that the volume of this research was to double every 10 years, but in 1990, it was reported that with development of information technology, Amount of information was 2x every 20 months. In such circumstances, the identification of subject areas and the relationship between them was difficult for researchers and professionals. Moreover, science and research policymakers experienced difficulties in identification of the dynamics of science and its research planning (He, 1999).

Library and information science (LIS) is no exception particularly that for the years, library and information science as a interdisciplinary and epistemic field has attracted a lot of attention and in recent years to keep pace with global developments in science and technology and in accordance with the needs of communities, new trends and courses have been derived from it. Thereby, identifying subject areas have attracted most attention and relation of those areas with together, is necessary.

The main purpose of this study is to investigate the LIS in the top countries of Middle East in term of scientific productions (Iran, Israel, Turkey, Kuwait, and Saudi Arabia), what structure and extent has changed during two decade 1992-2012.

In order to investigate the structure of each field, Innovative methods and different tools have emerged which one of the methods is analysis of article's title word that seems researchers have not paid meticulous attention (Milojevic, Sugimoto, Yan and Ding, 2011) while title words are appropriate for the aim given that "titles of journal articles themselves have undergone a change during the 20th century, becoming more informative, more specific, and containing a larger number of words that indicate article content" (Buxton and Meadows, 1977; Meadows, 1998 as cited in Milojevic´and et al, 2011). In according to Leydesdorff (1989) title words could be considered as tool of making visible the internal cognitive structure of a discipline. Thereby, it seems rational that authors identify cognitive structure and disciplinary diversity of knowledge area throughout title (Milojevic´, 2012) , in particular within the last two decades, this technique, implemented by various researchers, has been known as appropriate means for knowledge discovery in databases (He, 1999).

In study of fields, investigating the evolution of each fields or disciplines is important because the evolution indicates how interests in scientific society have changed over a long time. So, methods such as Co-word analysis developed. Co-word analysis is a techniques based on the co-occurrence frequency of pairs of words or phrases, this method is used to discover linkages among subjects in a research field and thus to trace the development of science (He, 1999; Muñoz-Leiva, Viedma-del-Jesús, Sánchez-Fernández, López-Herrera, 2012). Additionally, co-word analysis is utilized in a longitudinal framework which allows us to analyze and track the evolution of a research field along consecutive time periods (Garfield, 1994). As a finding of the method researchers are able to quantify and visualize the thematic evolution of the CBR (Consumer Behavior Research). Meanwhile, it helps to both experts and novices to predict where future research could lead (Muñoz-Leiva and et al, 2012).

Related works

In this work to present related work, we have used the Nazari's suggested pattern (2013). This pattern expresses analytically previous studies from subject and methodological view. Therefore, related works in respect of methodological lens are as follow:

Topic analysis (Sugimoto, Li, Russell, Finlay, and Ding, 2011; Sugimoto and McCain, 2010); analysis of word-reference combinations (van den Besselaar and Heimeriks, 2006); tri-occurrence analysis of index terms (Sugimoto and McCain, 2010); co-word analysis of both index terms and words extracted from titles, abstracts, and full text (Åström, 2002; Ding, Chowdhury, and Foo, 2001; Janssens, Leta, Glänzel, and De Moor, 2006); bibliometric analysis of authors (Moya-Anegón, Herrero-Solana and Jiménez-Contreras, 2006; White&McCain, 1998); bibliometric analysis of journals and journal articles (Åström, 2007, 2010; Moya-Anegón et al., 2006; Persson, 1994); content analysis (Järvelin and Vakkari, 1990, 1993).

However, in terms of subject lens, all of the researches limited to LIS, LS, IS and related fields.

Methodology

In this study we decide to investigate cognitive structure and evolution of Library and Information Science (LIS) in the top countries in the Middle East between 1992- 2012. For this aim, we use method of word analysis for title phrases from LIS research articles and select the articles are published in two decades (1992-2012) to study both the cognitive structure and evolution over this time period.

Data Collection

In order to gather the research's data we use several source. Firstly we selected five top countries using the S.J.R. website, (www.scimagojr.com) this site ranks the countries and journal according to scientific productions. Secondly, LIS articles related to each country retrieved from the Web of Science (WoS) 2013 database produced Thomson Reuters. For this purpose, we search only articles of LIS in English language in time span 1992 to 2012 using WoS' advance search section (May 20, 2014), in next step all of retrieved articles have been exported in Excel format and finally, the title column was chosen as an analysis source.

To analysis the title words, we utilize WordStat software which is content analysis software and it is applied for studying text information such as title of articles, journals and lectures and for classifying text while uses stemming method and maps hierarchical clusters of co-occurrence relationships. It identifies themes and trends as, serves as useful tool to extract and analysis a large number of document's information and it can be used for information and commercial purposes (WordStat, 2013). Given that WordStat cannot simultaneously calculate the phrases and the single words, so due to the more connotations of phrases in comparison with the single word they are chosen (Meadow, 2008). For this we considered minimum length term 2 and maximum length term 3 with maximum frequency 3.

Furthermore, VOSviewer software will be used for co-word analysis. The program can for instance be used to create maps of publications, authors, or journals based on a co-citation network or to create maps of keywords based on a co-occurrence network (VoSviewer, 2013).

Findings

Structure of LIS

In this research we investigated around 1000 articles using co-occurrence technique of title words to determine their hierarchical clustering and thus explain the internal structure of LIS. Given that word frequency is an important measure in content analysis and title words depict article contents, so it applies to identify the most important research topics or concepts in a field. These concepts are demonstrated as a dendrogram based on Jaccard's index, is shown in Figure1.

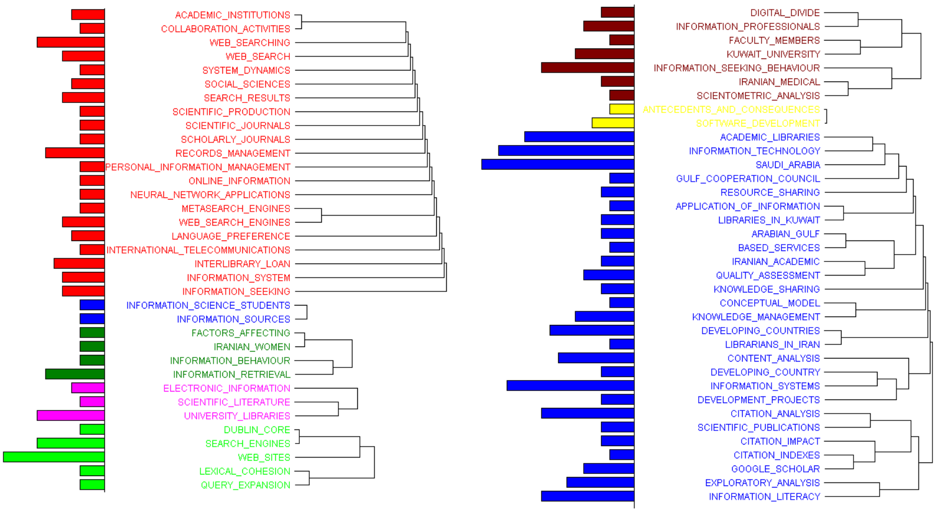


Figure 5. The dendrogram of the second-order clustering of whole of terms (phrases) based on Jaccard index

Note: whole of the figure is a dendrogram. For demonstration ease is divided into two part

This figure displays cognitive structure of LIS into 8 clusters seprated with different colour and we do not considerd clusters separately to analysis the content and the structure of LIS and dendrogram as whole has been considered. So, LIS structure are devided to three principal fields, as follow:

1. **Library Science** which comprises subfields such as *Theoretical Studies*, *Information Document, Social Science, Interlibrary Loan, Knowledge Management, Developing Countires, Resouce Sharing* and More importantly *Academic Libraries*. As can be seen in figure 2, *Collaboration, E-sources, Information Technology, Information Documents, Quality Assessment, Knowledge Sharing* have mainly been considered as the significant subfields in Academic Library.
2. **Information Science and Technology** which include three subfield "Information Science", "Web" and "Information Technology". Therefore, in "Information Science" Field *Information Mangement, Information Systems, Information Retrivel, Information Services, Information Literacy* have considerably been expressed. "Web" ,as another important subfield, is consisted of topics such as *Web Search, Search Enginge, Metadata, Web Site* while in "Information Technology", *Neural Networks, Computational Linguistics, Telecommunication and Software Development* have been taken into consideration as significant topics in this subfield.
3. **Bibliometric and Scientometric**, in according to interpreting Dendrogram structure *Citation Analysis, Citation Impact and Collaboration* are related topics to this subfield (See figure 2).

Figure 6. The branches of LIS which obtained of second-order clustering using the Jaccard index

Evolution of LIS

To identify how the LIS have changed in the two decades, we investigated the terms used in the article titles and for this we utilized the VoSviewer software that the changes have been presented in Figure 3, 4, 5 and 6. As can be seen, position of LIS in two periods has been depicted in the "label view" and "density view".

The circles in the "label view" map (figures 3 and 4) are representative the terms which their size and color specify the frequency and subject field respectively.

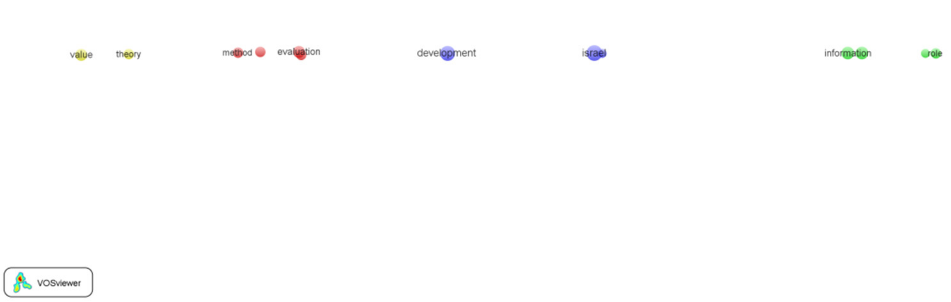


Figure 7. Label View of LIS during the 1992-2002 (decade 1)

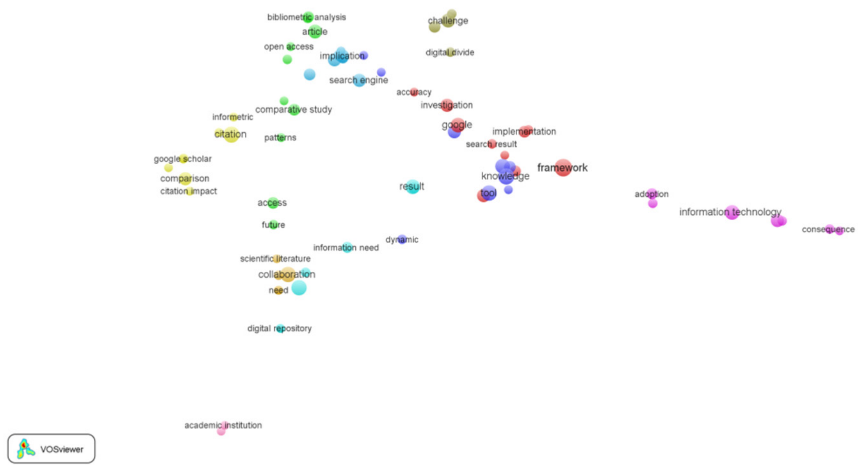


Figure 8. Label View of LIS during the 2003-2012 (decade 2)

As the maps illustrate, in the first decade the number of the LIS published papers is not considerable and have poor topic diversity and more importantly, relations among them has not been depicted well. But in the second decade the number of articles has been increased and hence the topic relations among the circles and fields have been represented well than previous decade.

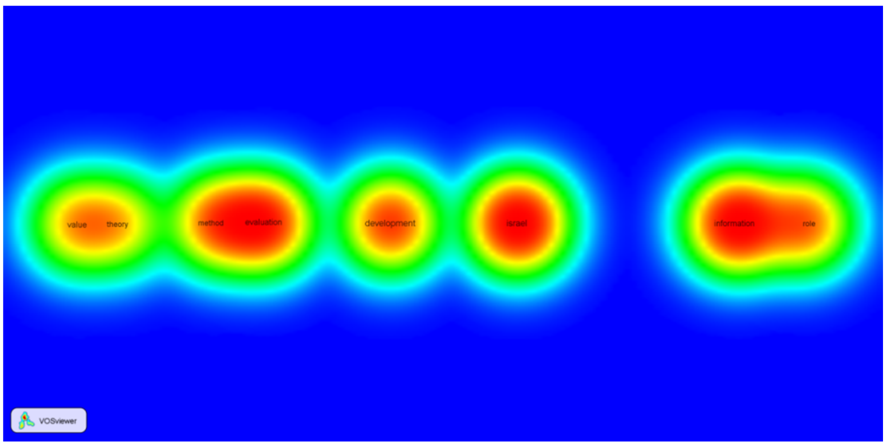


Figure 9. Density View of LIs during 1992-2002 (decade 1)

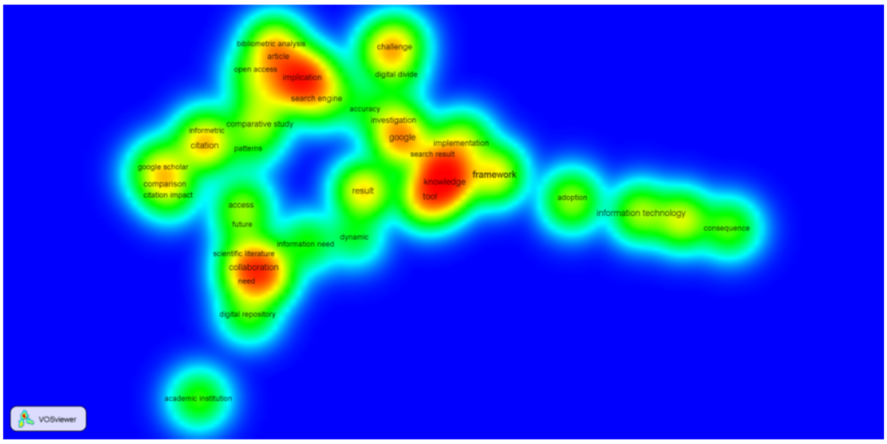


Figure 10. Density View of LIS during 2003-2012 (decade 2)

In according to the "density view" map (figure 5 and 6), the important regions and terms have more density are clearly visible. As the figures 5 and 6 show, by moving from red regions to blue regions the density decreases. In the other word, the important terms and subjects exist in the red and orange regions. In the first decade although the topics have not been expressed well but all terms of any clusters are hot and belong to red or orange regions. In contrary, in the second decade each cluster has a few hot terms; some clusters have no hot terms. The terms and hot terms of each cluster in the first and second decade have been demonstrated in the table 1 and 2.

Table 1. Term & hot term of LIS during 1992-2002 (decade 1)

Name of cluster	Terms in the each cluster	Hot terms in the each cluster
Cluster 1	Information, Role	Information, Role
Cluster 2	Development, Israel	Development, Israel
Cluster 3	Method, Evaluation	Method, Evaluation
Cluster 4	Value, Theory	Value, Theory

***Note**: clustering and terms in each cluster based on figure 3 and hot terms based on figure 5

Table 2. Term & hot term of LIS during 2003-2012 (decade 2)

Name of cluster	Terms in the each cluster	Hot terms in the each cluster
Cluster 1	Adoption, information technology, consequence	-
Cluster 2	Framework, implementation, search result, google, investigation, accuracy	Search result, google
Cluster 3	Knowledge, tools, dynamic	Knowledge, tools, dynamic
Cluster 4	Challenge, digital divide	Challenge
Cluster 5	Implication, search engine	Implication, search engine
Cluster 6	Bibliometric analysis, article, open access, patterns, comparative study, access, future	Article
Cluster 7	Informetric, citation, google scholar, comparison, citation impact	Citation, comparison
Cluster 8	Scientific literature, collaboration, need	Collaboration, need
Cluster 9	result, information need, digital repository	-
Cluster 10	Academic institute	-

***Note**: clustering and terms in each cluster based on figure 4 and hot terms based on figure 6

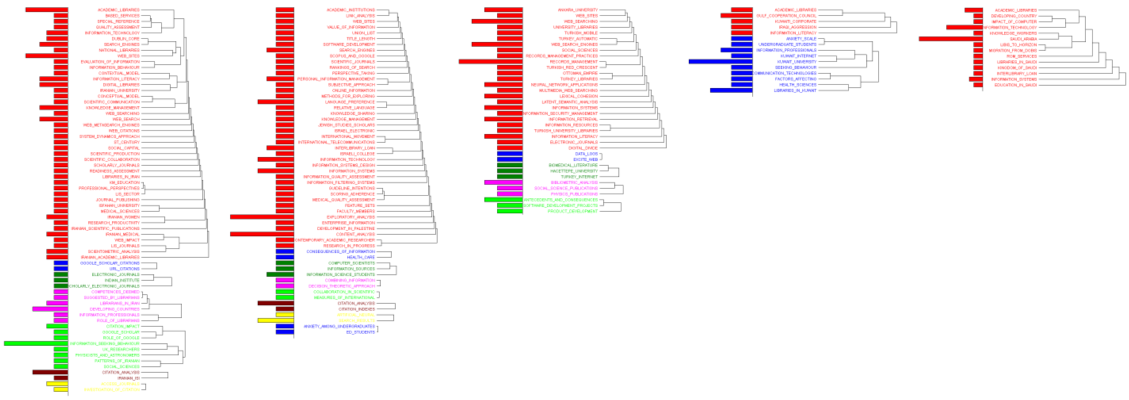


Figure 11. The dendrogram of the second-order clustering of whole of terms (phrases) based on Jaccard index/ Left to Right: Iran, Israel, Turkey, Kuwait, Saudi Arabia

Discussion and Conclusions

In this paper, using title phrases analysis and hierarchical clustering, we learned the structure of LIS over the two decades (1992-2012). Additionally, we attempt to present terms which have been used in each decade and evolution of LIS using the co-word analysis of LIS article titles.

The results show the structure of LIS in the top countries of Middle East in terms of scientific productions during the years of 1992-2012, in the several branches. These branches have been defined as Library Science, Information Science and Technology, and Bibliometric and Scientometric which has been got from whole of dendrogram and each of them include several subfield which some of the most important topics of each field as follow:

Knowledge Management, Interlibrary Loan, Academic Library, and Resource Sharing in the Library Science field.

Web Search, Search Engine, Web Site, Information Management, Information Retrieval, Information Seeking Behavior, Information Literacy, Information Technology in the Information Science and Technology and in the last field *Citation Analysis* is the unique topic.

Generally speaking, we could describe LIS structure and more importantly learn other useful findings in respect to dendrogram which seems essential to mention in this section. There are some main topics to which have not been paid attentions, such as Archives, Manuscripts, Cultural Heritage, National, Public and School Library.

Overall, LIS cognitive structure in Middle East during two decades reflects general tendency to Information Science and Technology field, in particular in Iran, Israel and Turkey. Also, topic diversity in the dendrogram of these countries sounds remarkably outstanding, particularly in Iran and Israel, as they have paid meticulous attention to Information Science and Technology and Bibliometric and Scientometric fields. Bearing in mind that Iran possesses a bold presence in Scientometric field given that recently Scientometric collage as a main branch of Library and Information Science major has been established. Furthermore, both Academic Libraries and Interlibrary loan among Arabian Gulf (Authors believe that Persian Gulf is accurate term!) Region countries are the striking features in the dendrogram of Kuwait and Saudi Arabia (see figure 7).

The findings about the evolution express that in the first decade Israel has a significant presence in comparison with other countries. Meanwhile, the number of released articles does not sound remarkable, so LIS cognitive structure definitely have not completely formed yet in this time span. The main concentration of the publications is on Theoretical Studies such as theories, values, roles, methods. In according to results in map 3, information term advent demonstrates its importance in 90s decade which had been nominated as the information or the information explosion age.

However, in the second decade we encounter with an entirely different scenario, as evolution trend of LIS is clearly visible. We can express that the main structure of LIS has been totally formed in this decade. Scientometric, Web and Information Science domains appearance, which constitute the main body of LIS, is obviously clear. Besides, Knowledge term has been interestingly substituted with Information given that 21 century has been identified as knowledge-based age. To pay serious attention to future studies (*future*), information documents analysis (*articles*) could be considered as the dominant evolution in LIS. Also, a part of publications have been carried out based on comparative approach (*comparative studies*). Due to Scientometric domain emergence in this decade, collaboration as the principal part in this study domain has been expressed. Furthermore, there are some terms with high frequency such as *framework* and *challenge* which only apply to formatting titles and do not belong to the particular domain.

Findings of our study show that these analyses are appropriate approach to demonstrate the structure and evolution not only of LIS but also for each other of scientific fields. Although, the different interpretation of dendrogram and maps could be existed and this depends on each researcher, however, this work is an important and useful step toward using title phrases and co-word analysis to identifying the structure and evolution of each discipline.

Acknowledgments

The authors express enough thanks to Encyclopaedia Islamica Foundation (EIF) ' chiefs for their continued encouragement.

References

Åström, F. (2002). Visualizing library and information science concept spaces through keyword and citation based maps and clusters. In H. Bruce, R. Fidel, P. Ingwersen, & P. Vakkari (Eds.), *Emerging frameworks and methods*. CoLIS4: Proceedings of the Fourth International Conference on Conceptions of Library and Information Science, July 21-25 (pp. 185–197). Greenwood Village, CO: Libraries Unlimited. Retrieved April 23, 2014 From http://lup.lub.lu.se/luur/download?func=downloadFile&recordOId=579259&fileOId=579260

Åström, F. (2007). Changes in the LIS research front: Time-sliced co-citation analysis of LIS journal articles, 1990–2004. *Journal of the American Society for Information Science and Technology*, 58 (7), 947–957.

Buxton, A.B., & Meadows, A.J. (1977). The variation in the information content of titles of research papers with time and discipline. *Journal of Documentation*, 33, 46–52.

Ding, Y., Chowdhury, G., & Foo, S. (2001). Bibliometric cartography of information retrieval research by using co-word analysis. *Information Processing & Management*, 37 (6), 817–842.

Garfield, E. (1994). Scientography: mapping the tracks of science. *Current Contents (print editions)*, 7, 5–10

He, Qin (1999). Knowledge Discovery Through Co-Word Analysis. *Library Trends*. 48 (1): 133-159. Retrieved February 1, 2013 from https://www.ideals.illinois.edu/bitstream/handle/2142/8267/?sequence=1

Janssens, F.; Leta, J.; Glänzel, W.; De Moor, B. (2006). Towards mapping library and information science. *Information Processing & Management*, 42 (6), 1614–1642.

Järvelin, K.; Vakkari, P. (1990). Content-analysis of research articles in library and information science. *Library & Information Science Research*, 12 (4), 395–421.

Järvelin, K.; Vakkari, P. (1993). The evolution of library and information science 1965–1985: A content analysis of journal articles. *Information Processing & Management*, 29 (1), 129–144.

Leydesdorff, L. (1989). Words and co-words as indicators of intellectual organization. *Research Policy*, 18, 209–223.

United Kingdom: Emerald. Meadow, Charles T. (2008). Text information retrieval systems

Meadows, A.J. (1998). *Communicating Research*. San Diego, CA: Academic Press.

Milojević, Staša (2012). Multidisciplinary Cognitive Content Of Nanoscience and Nanotechnology. *Journal of Nanoparticle Research*, 14 (1). Retrieved May 3, 2014 from http://arxiv.org/ftp/arxiv/papers/1204/1204.0127.pdf

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

Moya-Anegón, F., Herrero-Solana, V., & Jiménez-Contreras, E. (2006). A connectionist and multivariate approach to science maps: The SOM, clustering and MDS applied to library and information science research. *Journal of Information Science*, 32 (1), 63–77.

Muñoz-Leiva, Francisco; Viedma-del-Jesús, María; Sánchez-Fernández, Juan; López-Herrera, Antonio (2012). An application of co-word analysis and bibliometric maps for detecting the most highlighting themes in the consumer behavior research from a longitudinal perspective. *Quality & Quantity*, 46 (4), 1077-1095. Retrieved April 23, 2014 From http://sci2s.ugr.es/publications/ficheros/QuQu.pdf

Nazari, Maryam (2013). How knowledge gap is captured in generative research? a proposal for eveloping two maps: knowledge map and research map. *Journal of Academic Librarianship and Information Research*, 47 (1), 27-48.

Persson, O. (1994). The intellectual base and research fronts of JASIS 1986– 1990. *Journal of the American Society for Information Science*, 45 (1), 31–38.

Price, D. S. (1963). *Little science, big science*. New York: Columbia University Press.

Provalis research (2013). Retrieved from http://provalisresearch.com/products/content-analysis-software/

Sugimoto, C.R.; Li, D.; Russell, T.G.; Finlay, S.C.; Ding, Y. (2011). The shifting sands of disciplinary development: Analyzing North American Library and Information Science dissertations using Latent Dirichlet Allocation. *Journal of the American Society for Information Science and Technology*, 62 (1), 185–204.

Sugimoto, C.R.; McCain, K.W. (2010). Visualizing changes over time: A history of information retrieval through the lens of descriptor tri-occurrence mapping. *Journal of Information Science*, 36 (4), 481–493.

van den Besselaar, P.; Heimeriks, G. (2006). Mapping research topics using word-reference co-occurrences: A method and an exploratory case study. *Scientometrics*, 68 (3), 377–393.

VoSviewer (2013). Retrieved from http://www.vosviewer.com/

White, H.D.; McCain, K.W. (1998). Visualizing a discipline: An author co-citation analysis of information science, 1972–1995. *Journal of the American Society for Information Science*, 49 (4), 327–355.

Weight of Webometrics Criteria using Entropy Method

Handaru Jati

Universitas Negeri Yogyakarta Karangmalang Campus Indonesia
handaru@uny.ac.id

Abstract

The aim of this study was to propose some of the basic tools for Decision Making. The purpose of this paper is to show a methodology test for the selection of the weighted method, as aid to decision making in the design stage in the area of webometrics. Selecting the weighted method is one of the problems of Multicriteria Decision Analysis in which decision-makers have had disadvantages in weighting assignment criteria. To resolve this problem arises weighting variables using the entropy method. The model presented in this article is limited to display application in a webometric case. This model can be applied as a way to supplement the technical studies to select the weighted method of a webometrics and it gives the relative importance weights of the various elements, and gives an empirical analysis, explain the role of the entropy weight in webometrics study. Entropy weighted method enables rank all the alternatives in question without decisor bias and calculates the specific weight of criteria.

Introduction

In the world there are thousands of universities, and since 2004 it has been published a Web Ranking whichshows the results in every six months (January and July) and covers about 20,000 Higher Education Institutions worldwide. The composite index (Ranking) is calculated by combining standardized positions instead of values. The visibility is calculated giving an extra inbound links that are not from generic domain importance (.Com, .Org, .Net). Figures for rich files (pdf, doc, ppt, ps, Dox, pptx, eps) are combined and have not been treated individually. The intention with this system of analysis and projection of cybermetric indicators under the parameters set Webometrics is to strengthen and indicate the type of information being generated in each of the institutions and thereby improve certain characteristics that further enrich university of university webometrics ranking has changed the setting of higher education and is likely to continue to influence further development nationally and internationally. This moment is a new era for university, characterized by global competition, in which university ranking systems have assumed an importance factor for surviving. Their emergence *has* also *been* a matter of *controversy*, often controversial and subject to considerable debate, has been met with a lot of scepticism, some enthusiasm and an institutional unease. Academic rankings are here to stay and it is results that count for most of higher education's stakeholders.

Literature Review

Webometrics

Although the subfield of webometrics is considered as one of the most recent quantitative studies within the field of library and information science, there are already several international studies that address this topic. Many authors have directed their focus of study for this new environment, for finding web immense diversified network of information resources, easily accessible and still little explored. In this sense, Cronin and McKim (1996) argue that as the Web is becoming a medium increasingly important to science and academia , it is logical that quantitative studies extend well to this medium. Also Thelwall, Vaughan and Björneborn (2003) consider that being a global network of Web documents, initially developed for academic use and then extended to general users , it is obvious that it is a fertile field of research for bibliometrics , the scientometrics and informetrics. The Webometrics is a ranking based on measurements of the presence of the universities on the Web. It is prepared by the Laboratory Cybermetrics, a group of research is part of the Superior Council of Scientific Research of Spain, and not for commercial purposes. In contrast to other rankings, Webometrics classifies a large number of universities, more than 20,000 in its latest edition (January 2012). Published twice a year (January and July). The system also allows universities ordered by country and region (Aguillo, Ortega et al. 2008). According to its website, the ranking aims to promote open access to information on the Internet by universities access. Also, as most of the rankings, insist on the superiority of his method: "As other rankings focused only on a few relevant aspects, specially research results, our ranking based on indicators of the presence reflects best the Web overall activity of the institutions, as there are many other tasks performed by teachers and researchers that appear on the Web. However, this method also has its limitations, since it favors large universities or those with large budgets for technology.

Entropy Method

The entropy method was developed as an objective method of allocation weights depending on the decision matrix without affecting the preference of the decision maker (Zeleny 1982), the relative importance of criterion j in a decision situation, wj measure its weight is directly related to the amount of information provided by the intrinsically set of alternatives with respect to that criterion (Barba Romero and Pomerol 1997). How much have greater diversity in the evaluations of the alternatives greater importance should be the criterion. Far this diversity is conceptually based on solid and accepted concept of entropy in an information channel posed by Claude Shannon (Shannon and WEAVER 1949) . The procedure is as follows:

a. The evaluations ij (i = 1, m) (j = 1, n) are taken as normalized as a fraction of the sum i ij Σ to the original assessments of each criterion j.

$$a_{ij} = \frac{k_{ij}}{\sum_{i=1}^{m} \sum_{j=1}^{n} k_{ij}} \quad \text{for m > 1 and i=1, 2, …, m; and j=1, 2, …, n.} \tag{1}$$

b. Entropy (Ej) is calculated.

$$E_j = [\frac{-1}{\ln(m)}]\sum_{i=1}^{m}[a_{ij}\ln(a_{ij})] \tag{2}$$

where m = number of alternatives in the matrix standardized assessments and ij = Criteria or standardized attributes.

c. Diversity criterion (Dj) is calculated.

$$D_j = 1 - E_j \tag{3}$$

d. The normalized weight of each criterion (Wj) is calculated.

$$w_j = \frac{D_j}{\sum D_j} \tag{4}$$

Research Method

Weighted indicators that take into account are:

- Size: number of pages recovered from 4 search engines: Google, Yahoo, Live Search and Exalead (20%).
- Visibility: The total number of unique external links received (inlinks) by a site that you can den get consistently from Yahoo Search, Live Search and Exalead (50%).
- Rich files: the following file formats were selected after considering their relevance in academic and publication activities and considering the volume of use: Adobe Acrobat (pdf.), Adobe PostScript (ps.), Microsoft Word (. Doc) and Microsoft Powerpoint (.

Ppt). These data are extracted through Google, Yahoo Search, Live Search and Exalead (15%).

- Academic: Google Scholar provides the number of papers and citations for each domain academic. The results obtained from the database of Google Scholar papers, reports and other academic papers (15%).

Results

The four number of criteria that should typically be considered in selecting the best university website are Size(C1), Visibility (C2), Rich Files (C3), and scholar (C4). First of all we form the decision matrix, after that we compute hi ,di and wi base on Shannon method that are shown in Table 1.

Table 1. Data

Universitas	Size	Visibility	Rich Files					Scholar
			.pdf	.ps	.ppt	.doc	Total	
Uni A	9950	177,321	259000	84200	9110	22900	375210	9950
Uni B	8970	307,113	390000	26400	10800	13400	440600	8970
Uni C	33200	4.616,437	317000	22300	18100	19900	377300	33200
Uni D	30100	362,854	268000	10100	8650	22800	309550	30100
Uni E	26700	113,286	269000	12900	20000	20500	322400	26700

We want to obtain a weight for each criterion by using the proposed approach. According to Eq.1, normalized matrix data are presented.

Table 2. Normalized Data

Size	Visibility	Rich Files	Scholar
0,040	0,014	0,501	0,040
0,000	0,043	1,000	0,000
1,000	1,000	0,517	1,000
0,872	0,055	0,000	0,872
0,732	0,000	0,098	0,732

The evaluations of these five alternatives according to the previously stated criteria, i.e., evaluation matrix, are displayed in Table 3.

Table 3. Normalized Data

Size	Visibility	Rich Files	Scholar
0,960	0,986	0,499	0,960
1,000	0,957	0,000	1,000
0,000	0,000	0,483	0,000
0,128	0,945	1,000	0,128
0,268	1,000	0,902	0,268

In our analysis we calculate diversity criteria and the result shows in the table 4.

Table 4. Diversity Criterion

	Size	Visibility	Rich File	Scholar
	-0,665112338	-0,683290598	-0,345859551	-0,665112338
	-0,693147181	-0,663312374	0	-0,693147181
	0	0	-0,334805162	0
	-0,088681645	-0,654732445	-0,693147181	-0,088681645
	-0,185945385	-0,693147181	-0,625181204	-0,185945385
E(C) = ln(2) *total sum	**0,471151465**	**0,777463336**	**0,576787486**	**0,471151465**
d = 1-E(C)	**0,528849**	**0,222537**	**0,423213**	**0,528849**

The final rank of each criterion by using the entropy weighted method can be seen in table 5. The obtained values of criterion Size, visibility, rich files, and scholar are 0,310458; 0,130639; 0,248445; and 0,310458 respectively.

We see that the rank of size and scholar are just better than the rank of rich file and visibility. Therefore, size locates at rank 1. Other criteria can be ranked in the same way. For problems with more complexity, with a small program (for example Excel) we can determine the rank of each criterion. In the last Table 5, the rank of each criterion can be seen.

Table 5. Weight of Criterion

Criteria	Weight (W) = d/total
Size	**0,310458**
Visibility	**0,130639**
Rich File	**0,248445**
Scholar	**0,310458**

Conclusion

There are several methods for obtaining the weights of criteria of an MADM problem, one of which is the entropy method.How to ascertain weights and subjectivity of evaluation model are the main aspects which influence evaluation result in the present quantitative evaluation methods. During ascertaining weights, either subjectivity can't be avoided, or calculation is too complex. On the other hand, subjectivity can't be avoided in some evaluation methods. based on entropy weight can avoid not only subjectivity or complex calculation in ascertaining weights but also subjectivity of evaluation model via the evaluation criteria of weighted relative adjacent degree. Entropy weighted method is a new advancement in quantitative evaluation methods for webometrics.

References

Aguillo, I. F., J. L. Ortega, et al. (2008). "Webometric ranking of world universities: Introduction, methodology, and future developments." *Higher Education in Europe* 33: 233-244.

Barba Romero, S. and J. Pomerol (1997). Decisiones Multicriterio Fundamentos Teóricos y Utilización Práctica. *Universidad de Alcalá de Henares*.

Shannon, C. E. and W. WEAVER (1949). *The Mathematical Theory of Communication*. Chicago, University of Illinois Press.

Zeleny, M. (1982). *Multiple Criteria Decision Making*. New York, Mc Graw Hill.

Detecting the milestones of epigenetics development from 2002 to 2013: a Scientometrics perspective

Hongfang Shao*, Qi Yu** and Zhiguang Duan***

*School of Public Health, Shanxi Medical University. No.56, South Xinjian Road, Taiyuan 030001, China
shftanya@126.com

**Department of Information Management, Shanxi Medical University.
No.56, South Xinjian Road, Taiyuan 030001, China
yuqi351@gmail.com

***School of Public Health, Shanxi Medical University. No.56, South Xinjian Road, Taiyuan 030001, China
dzg52827@aliyun.com

Abstract

The study aims at creating a picture of how epigenetics develops and evolves over the past decade and detecting its intellectual milestones. The documents which contain the word 'epigenetic' or 'epigenetics' in their title, abstract or keywords were retrieved from Web of Science between 2002 and 2013 to be our dataset. Co-citation analysis, betweenness centrality, citation bursts, and the newest scientometrics methods were used in this study.

By mapping the discipline co-citation of epigenetic research over the last decade, we identified the major disciplines involved in epigenetics and found DNA methylation and histone modification are mainly the epigenetic research hotspots over the past 10 years, ncRNA is one main research front over the last 5 years. Meanwhile, we detected top 40 significant milestone literatures over the last decade. Our results show epigenetics has become more and more feasible in many disciplines and that it will possibly revolutionize clinical and healthcare practice and many aspects of our biomedicine. Visual analytics of the literature provides a valuable, timely, repeatable and flexible approach so as to track the epigenetics development process, which we expect will play a more active role in supplement to traditional survey articles.

Introduction

Epigenetics, a subdiscipline of genetics studying epigenetic variation, is a rapidly growing and fast-moving interdisciplinary field of study, involving the genetic gene expression changes without DNA sequence changes (Goldberg et al., 2007; Morris and Wu, 2001). In information age, as research in these areas advances rapidly, how to quickly understand the resolved and unresolved issues, proven or unproven methods, and to search for new discoveries, groundbreaking theories and technology is critical for scientific researchers across vast amounts of literature data.

Science knowledge mapping tools typically take scientific literatures as an input and generate visual mappings of complex structures for statistical analysis and multidimensional visualization exploration (Small,1973; Cobo et al., 2011). An array of science mapping tools are made widely available to researchers and analysts, notably including HistCite (Garfield, 2004), VOSviewer (van Eck and Waltman, 2010), Network WorkBench (Borner et al., 2010), DIVA (Morris et al.,2003), Loet Leydesdorff's software (Leydesdorff and Schank,2008) and CiteSpace (Chen et al.,2004,2006, 2008,2010,2012), and in biomedicine research, such as Cytoscape (Shannon et al., 2003) and Ingenuity Pathway Analysis (Jimenez et al., 2009).

CiteSpace is specifically designed to facilitate the detection of emerging trends and abrupt changes in scientific literature (Chen et al., 2012). In this study, we demonstrate a scientometric approach and use CiteSpace to delineate the structure and dynamics of the epigenetic research and detect groundbreaking intellectual milestones between 2002 and 2013.

Materials and Methods

Materials

The documents which contain the word 'epigenetic' or 'epigenetics' in their title, abstract or keywords were collected from the scientific literature database of Web of Science from 2002 to 2013. Only documents type of article and review were taken into account, filtering out proceedings papers and notes, the dataset was reduced to 27,117 records. We expanded the dataset by citation indexing. The citation index-based expansion resulted in 125,204 records. Thus, the 125,204--article dataset is used in the subsequent analysis.

Methods

Co-citation analysis

Co-citation analyses include document co-citation analysis, author co-citation and discipline co-citation analysis, etc. Here, we introduce document co-citation. The principle is as follows: if two documents are cited together in one or more articles, we say that the two documents are co-cited. Higher co-citation frequencies indicate closer links between the documents. Thus, based on the document citation relationship, we can analyse the affiliations between documents. If the documents are divided into clusters and classes, we can analyze the fronts of the current research according to the contents of documents (Chen et al., 2010; Liu et al., 2008).

Betweenness centrality

The betweenness centrality of a node in the network measures the importance of the position of the node in the network. It measures the extent to which the node is in the middle of a path that connects other nodes in the network. The betweenness centrality node called hub node shows where a particular article has connections to many different papers. A widely co-cited hub article is a good candidate for significant intellectual contributions (Chen et al., 2006, 2010).

Citation bursts

Burst detection determines whether a given frequency function has statistically significant fluctuations during a short time interval within the overall time period. It is valuable for citation analysts to detect whether and when the citation count of a particular reference has surged. It can be also used to detect whether a particular connection has been significantly strengthened within a short period of time. The approach identifies the "burstiest" articles and the time interval when they were active, to generate a timeline and follow the hotspots in a field (Chen et al., 2006, 2010).

Geometric mean

The geometric mean of the study refers to the geomean metric of betweenness centrality and citation bursts. The geometric mean measures both structural centrality and citation burstness of a cited reference. If a reference is strong in both measures, it will have a higher geometric mean than a reference that is only strong in one of the two measures（Chen, 2011）.

Results and Analysis

Science knowledge maps (Fig. 1 and Fig. 3) were accomplished automatically by CiteSpace II. In order to more accurately identify the intellectual milestones of epigenetic development, we divided the years "2002-2013" into four 3-year slices to detect the development track of epigenetic (Fig. 1). For example, the caption"2002-2004, N=230, E=810. 2,2,20"left- above the first snapshot of the network means that the network was formed between 2002 and 2004, consisting of 230 references and 810 co-citation pairs. Each reference has received at least 2 citations in one of the 3 years during this period.

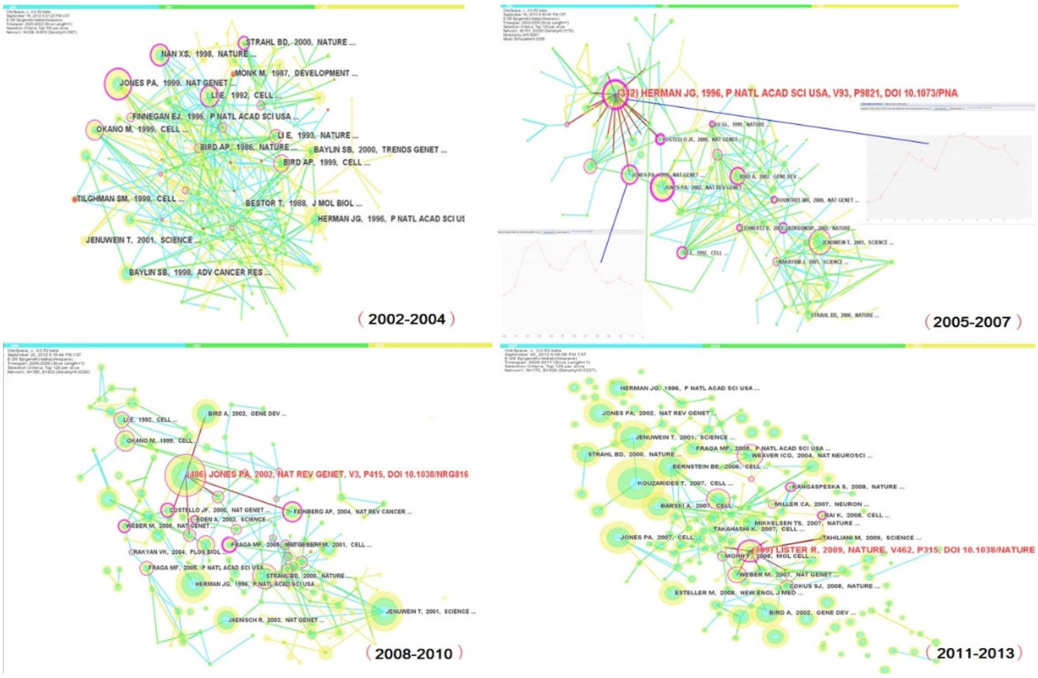


Figure 1. The four 3-year slices of epigenetic development snapshots

Fig.1 shows four 3-year snapshots of the co-citation network as it evolved over time. In each diagram, three colors match to the 3 years in the order of cyan, green and yellow. Thus, a yellow cluster would be formed in the 3rd year of a given 3-year interval. And, a node with essentially a yellow tree-ring means the reference was mostly cited in the 3rd year of the time interval. The captions left-above network snapshots record the time interval, the number of nodes, the number of co-citation links.

In the four snapshots, there are some important references continually detected, such as：Jones PA-1999(Jones and Laird,1999), a review of epigenetic mechanism of cancer proposed the role of DNA methylation in gene inactivation in cancer would become a hotspot; Herman JG-1996(Herman et al.,1996), the original article developed a novel method-Methylation-specific PCR (MSP), which founded the technology basis as it evolved over time. They both are detected to have considerable degrees of citation burst (Table 3).

Fig.2 shows a panorama view of epigenetics research in the entire time interval of the dataset (2002-2013). We analyze the intellectual structure of epigenetics research from four aspects below so as to detect milestones articles.

Most cited articles

Highly cited articles are usually regarded as the landmarks due to their sustained attention by scientists. In the diagrams of our study, the biggest citation rings represent the most cited articles.

Table 1 lists top 20 most cited articles over the last decade.

In Fig.2 and Table 1 the most cited article in our dataset is Jenuwein T-2001 (Jenuwein and Allis, 2001) with 1350 citations, followed by Herman JG-1996 (Herman et al., 1996) with 1258 citations. The third one is a review article by Jones PA-2002(Jones and Baylin, 2002) with 1220 citations. In Fig.2 the first three articles are respectively from C_{25} (histone modification), C_{11} (DNA methylation) and C_{52} (human carcinoma). Articles at the fourth to sixth positions are from C11 and C25, namely Bird A-2002 (Bird A, 2002), Kouzarides T-2007(Kouzarides, 2007) and Strahl BD-2000(Strahl and Allis, 2000). Bird A-2002 is from C11 (DNA methylation), and the fifth and the sixth's authors both have inspired intense interest in histone modification.

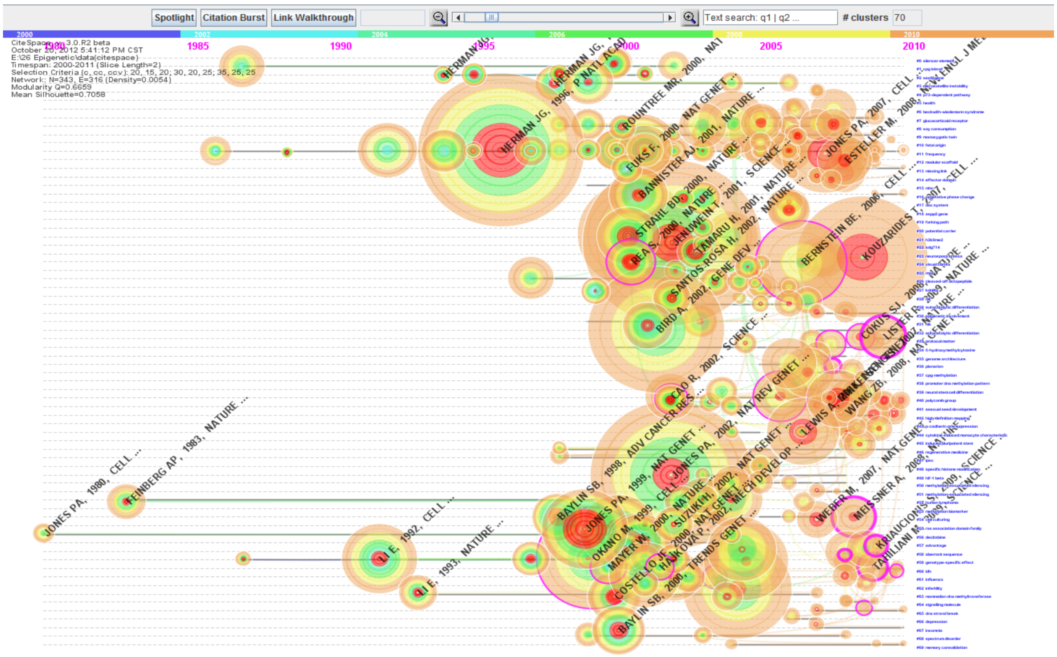


Figure 2. The panorama of epigenetics intellectual structure between 2002 and 2013. Landmark articles are labeled.

CiteSpace characterizes emerging trends and patterns of change in such networks in terms of a variety of visual attributes. The size of a node indicates how many citations the associated reference received. Each node is depicted with a series of citation tree-rings across the series of time slices. The structural properties of a node are displayed in terms of a purple ring. The thickness of the purple ring indicates the degree of its betweenness centrality. Such nodes tend to bridge different stages of the development of a scientific field. Citation rings in red indicate the time slices in which citation bursts, or abrupt increases of citations, are detected. Citation bursts provide a useful means to trace the development of research focus, or an emerging trend.

Table 1. Most cited references of top 20 (2002-2013)

No.	*Citation counts*	*References*	*burst*	*centrality*
1	1350	**Jenuwein T, Translating the histone code, 2001, SCIENCE, V293, P1074**	26.63	0.03
2	1258	**Herman JG, 1996, Methylation-specific PCR: A novel PCR assay for methylation status of CpG islands, P NATL ACAD SCI USA, V93, P9821**	103.95	0
3	1220	***Jones PA, 2002, The fundamental role of epigenetic events in cancer, NAT REV GENET, V3, P415**	63.09	0.04
4	1033	***Bird A, 2002, DNA methylation patterns and epigenetic memory, DNA methylation patterns and epigenetic memory, GENE DEV, V16, P6**		0.01
5	970	***Kouzarides T, 2007, Chromatin modifications and their function, CELL, V128, P693**	31.01	0.01
6	908	***Strahl BD, 2000, The language of covalent histone modifications, NATURE, V403, P41**		0.04
7	836	**Jaenisch R, 2003, Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals ,NAT GENET, V33, P245**	31.94	0
8	834	**Okano M, 1999, DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development, CELL, V99, P247**	27.94	0.11
9	691	***Jones PA, 1999, Cancer epigenetics comes of age, NAT GENET, V21, P163**	119.28	0.06

10	678	**Bernstein BE，2006，A bivalent chromatin structure marks key developmental genes in embryonic stem cells，CELL，V125，P315**		0.15
11	667	**Herman JG，2003，Mechanisms of disease: Gene silencing in cancer in association with promoter hypermethylation, NEW ENGL J MED，V349，P2042**	23.72	0
12	631	***Jones PA，2007，The epigenomics of cancer，CELL，V128，P683**	11.72	0
13	623	**LI E，1992，Targeted mutation of the DNA methyltransferase gene results in embryonic lethality，CELL，V69，P915**	49.68	0.01
14	609	**Cameron EE,1999,Synergy of demethylation and histone deacetylase inhibition in the re-expression of genes silenced in cancer,NAT GENET,V21,P103**	40.39	0.04
15	595	**Esteller M，2001，A gene hypermethylation profile of human cancer，CANCER RES，V61，P3225**	61.74	0
16	551	***Barski A，2007，High-resolution profiling of histone methylations in the human genome，CELL，V129，P823**	22.63	0.03
17	550	**Reik W,2001,Epigenetic reprogramming in mammalian development,SCIENCE,V293,P1089**	5.15	0.02
18	533	**Weaver ICG，2004，Epigenetic programming by maternal behavior,NAT NEUROSCI，V7，P847**		0
19	521	**Nan XS，1998，Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex，NATURE，V393，P386**	78.37	0.05
20	521	**Lachner M,2001,Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins，NATURE,V410,P116**	49.55	0.03

In 1996, Herman JG et al. described a new method, methylation-specific PCR (MSP), MSP eliminates the false positive results inherent to previous PCR-based approaches. This article is inevitably one of the milestones as a technique landmark. In 2002, Bird A et al. reviewed DNA methylation patterns and epigenetic memory (Bird A, 2002).

Jenuwein and Strahl et al. made great contributions to histone modification theory development for the proposition of the histone code hypothesis. In 2000, Strahl et al. first proposed the term“histone code”and presented a “histone code hypothesis” (Strahl and Allis,2000).

Betweenness centrality

The betweenness centrality of a node in the network measures the importance of the position of the node in the network. We are particularly interested in the method because it is likely to lead to insights into intellectual milestones in some scientific field.

Table 2 lists the top 20 structurally essential articles in the synthesized network between 2002 and 2013. These references are important in terms of not only how they connect individual nodes in the network but also how they connect aggregated groups of nodes, such as co-citation clusters. These works can be seen as landmark works in the context of our broadly defined area of epigenetic.

Table 2. Cited citations with the highest betweenness centrality of top 20 (2002-2013)

No.	*centrality*	*References*	*citations*	*burst*
1	0.31	**Lister R，2009，Human DNA methylomes at base resolution show widespread epigenomic differences，NATURE，V462，P315**	309	93.54
2	0.28	**Ball MP，2009，CTargeted and genome-scale strategies reveal gene-body methylation signatures in human cells,NATURE BIOTECHNOLOGY,V27,P361**	99	18.97
3	0.28	**Tahiliani M，2009，Conversion of 5-Methylcytosine to 5-Hydroxymethylcytosine in Mammalian DNA by MLL Partner TET1 ,SCIENCE，V324，P930**	202	55.92
4	0.27	**Kriaucionis S,2009,The Nuclear DNA Base 5-Hydroxymethylcytosine Is Present in Purkinje Neurons and the Brain,SCIENCE,V324，P929**	145	39.93
5	0.26	**Ramsahoye BH，2000，Non-CpG methylation is prevalent in embryonic stem cells and may be mediated by DNA methyltransferase 3a,P NATL ACAD SCI USA，V97，P5237**	160	6.63

6	0.22	**Meissner A，2008，Genome-scale DNA methylation maps of pluripotent and differentiated cells,NATURE，V454，P766**	310	45.73
7	0.18	**Popp C，2010，Genome-wide erasure of DNA methylation in mouse primordial germ cells is affected by AID deficiency,NATURE，V463,P1101**	97	29.47
8	0.17	**Mayer W,2000,Embryogenesis- Demethylation of the zygotic paternal genome,NATURE，V403,P501**	326	14.32
9	0.15	**Bernstein BE，2006，A bivalent chromatin structure marks key developmental genes in embryonic stem cells，CELL，V125，P315**	678	
10	0.13	**Rea S，2000，Regulation of chromatin structure by site-specific histone H3 methyltransferases，Nature,V406,P593**	403	46.3
11	0.13	**Muller J,2002,Histone methyltransferase activity of a Drosophila polycomb group repressor complex,CELL,V111,P197**	290	21.87
12	0.13	**Hajkova P,2002,Epigenetic reprogramming in mouse primordial germ cells,MECH DEVELOP,V117,P15**	243	
13	0.13	**Kaneda M,2004,Essential role for de novo DNA methyltransferase Dnmt3a in paternal and maternal imprinting,NATURE,V429,P900**	210	
14	0.11	**Okano M，1999，DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development，CELL，V99，P247**	834	27.94
15	0.11	**Boyer LA,2006,Polycomb complexes repress developmental regulators in murine embryonic stem cells,NATURE,V441,P349**	398	13.37
16	0.11	**Ohm JE,2007,A stem cell-like chromatin pattern may predispose tumor suppressor genes to DNA hypermethylation and heritable silencing,NAT GENET,V39,P237**	237	12.57
17	0.11	**Cokus SJ，2008，Shotgun bisulphite sequencing of the Arabidopsis genome reveals DNA methylation patterning，NATURE，V452，P215**	205	20.65
18	0.11	**Ooi SKT,2008,The colorful history of active DNA demethylation,CELL,V133,P1145**	117	18.61
19	0.1	**Widschwendter M,2007,Epigenetic stem cell signature in cancer,NAT GENET,V39,P157**	214	10.85
20	0.1	**Chan SWL,2005,Gardening the genome: DNA methylation in Arabidopsis thaliana,NAT REV GENET,V6,P351**	164	14.31

See Fig.2 and Table 2, Lister-2009(Lister et al.,2009) ranks 1st and comes from C52 with a prominent structural property—the highest betweenness centrality value of 0.31(a larger purple ring), which presented the first genome-wide, single-base-resolution maps of methylated cytosines in a mammalian genome. Ball-2009(Ball et al.,2009) may be a milestone paper of technology with the betweenness centrality value of 0.28, which ranks 2nd in Table 2 and comes from C11 and introduced two complementary high-throughput methylation profiling approaches. The third and fourth salient works are Tahiliani-2009 (Tahiliani et al., 2009) from C59 and Kriaucionis-2009 (Kriaucionis and Heintz, 2009) from C60.

Citation bursts

The automatically generated clusters include not only a highly cited article, but also with a strong surge of citations. A citation burst has two attributes: the intensity of the burst and how long the burst status lasts (Chen et al.,2012). It is critical for us to detect emerging trends and critical turns of the development of the collective knowledge.

Table 3 lists the strongest citation bursts articles of top 20 in the synthesized network between 2002 and 2013.

Top three articles with strong citation bursts are from C52 (human carcinoma) and C11(DNA methylation) and the fourth and the fifth are from C25 (histone modification) (Table 3 and

Table 3. References with the strongest citation bursts of top 20 (2002-2013)

No.	*Citation bursts*	*References*	*citations*	*centrality*
1	119.28	*Jones PA, 1999, Cancer epigenetics comes of age, NAT GENET, V21, P163	691	0.06
2	103.95	Herman JG, 1996, Methylation-specific PCR: A novel PCR assay for methylation status of CpG islands, P NATL ACAD SCI USA, V93, P9821	1258	0
3	96.23	Baylin SB, 1998, Alterations in DNA methylation: a fundamental aspect of neoplasia, ADV CANCER RES, V72, P141	449	0.06
4	93.54	Lister R,2009,Human DNA methylomes at base resolution show widespread epigenomic differences,NATURE,V462, P315	309	0.28
5	78.37	Nan XS, 1998, Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex, NATURE, V393, P386	521	0.05
6	73.64	*Baylin SB, 2000, DNA hypermethylation in tumorigenesis, TRENDS GENET, V16 , P168	452	0
7	73.32	Dammann R, 2000, Epigenetic inactivation of a RAS association domain family protein from the lung tumour suppressor locus 3p21.3, NAT GENET, V25, P315	421	0
8	72.84	Burbee DG, 2001, Epigenetic inactivation of RASSF14 in lung and breast cancers and malignant phenotype suppression,J NATL CANCER I, V93, P691	322	0
9	66.81	Tamaru H, 2001, A histone H3 methyltransferase controls DNA methylation in Neurospora crassa,NATURE, V414, P277	327	0.01
10	63.09	*Jones PA, 2002, The fundamental role of epigenetic events in cancer,NAT REV GENET, V3, P415	1220	0.04
11	61.74	Esteller M, 2001, A gene hypermethylation profile of human cancer, CANCER RES , V61, P3225	595	0
12	61.4	*Bird AP, 1999, Methylation-induced repression - Belts, braces, and chromatin, CELL, V99, P451	432	0
13	56.08	Bannister AJ, 2001, Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain, NATURE, V410, P120	481	0.05
14	55.92	Tahiliani M, 2009, Conversion of 5-Methylcytosine to 5-Hydroxymethylcytosine in Mammalian DNA by MLL Partner TET1 ,SCIENCE, V324, P930	202	0.28
15	54.56	*Esteller M, 2008, Molecular origins of cancer: Epigenetics in cancer, NEW ENGL J MED, V358, P1148	405	0
16	50.46	Herman JG, 1998, Incidence and functional consequences of hMLH1 promoter hypermethylation in colorectal carcinoma, P NATL ACAD SCI USA, V95, P6870	333	0
17	49.68	LI E, 1992, Targeted mutation of the DNA methyltransferase gene results in embryonic lethality, CELL, V69, P915	623	0.01
18	49.55	Lachner M,2001,Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins, NATURE,V410,P116	521	0.03
19	46.3	Rea S, 2000, Regulation of chromatin structure by site-specific histone H3 methyltransferases, Nature,V406,P593	403	0.13
20	45.73	Meissner A, 2008, Genome-scale DNA methylation maps of pluripotent and differentiated cells,NATURE, V454, P766	310	0.22

Fig.2). Jones PA-1999 review (Jones and Laird,1999) ranks 1st, Herman JG-1996 (Herman et al., 1996) ranks 2nd, Baylin SB-1998（Baylin,1998） ranks 3rd, Lister R-2009(Lister et al.,2009) and Nan XS-1998(Nan et al.,1998) rank 4th and 5th. Among these, Herman JG-1996 ranks 2nd in the sequence of the most cited articles (Table 1) and Lister R-2009 ranks 1st in the sequence of betweenness centrality (Table 2).

Geometric mean

For the total 60 articles listed in Table 1, 2 and 3, we found 11 articles appear twice, the other 38 appear once. From the citation frequency, citation burst and betweenness centrality analysis, we also found the three does not exist the consistency. For example, Strahl BD -2000(Strahl and Allis, 2000) has a strong citation frequency ranks 6th in Table 1 (with the citation of 908), but a lower burst of 8.02 and centrality of 0.04. Strahl BD-2000 belongs to the 38 articles, but not to the 11 ones, i.e. it appears once not twice; Lister R-2009(Lister et al., 2009) ranks 1st in Table 2 with centrality of 0.31 and ranks 4th in Table 3 with burst of 93.54, but a lower citation count of 309, which isn't listed in Table 1—The reason is there may be a great correlation between citation frequency and time: the earlier publications should relatively have higher frequency than the latter publications. The two articles are undoubtedly the landmark works, for one proposed the "histone code hypothesis", the other presented the first genome-wide, single-base-resolution maps of methylated cytosines in a mammalian genome. So, a number of questions can be addressed from our theory of discovery. Are the 11 articles of appearing twice more important than the 38 of appearing once listed in Table 1, 2 and 3? If we detect the milestones by most cited articles analysis, could we omit some important ones like Lister-2009? If we detect the landmark articles by centrality or by burst analysis, could we omit some ones like Strahl BD-2000?

With these questions, we went back to observe Table 1, 2 and 3 again, and found one literature had a higher citation frequency, whose burst and centrality value are not necessarily simultaneously high; but if a reference is strong in both burst and centrality value, it has a higher citation frequency than a reference that is only strong in one of the two. So we suggest considering the geometric mean of the burst and the centrality value as a comprehensive analysis of the influence of scientific literatures, which may likely be more reasonable. That is, the geometric mean in our study measures both structural centrality and citation burstness of a cited reference. If a reference is strong in both measures, it will have a higher geometric mean than a reference that is only strong in one of the two measures.

According to the geometric mean of betweenness centrality and citation burstness, we detected top 40 significant literatures in epigenetics research process. As shown in Fig.2, although we set up our original database between 2002 and 2013, the earliest milestone papers appear in 1980s because of the property of citation (Chen, 2004). From Fig.2 and Table 4, we conclude epigenetics research has definitely developed over time, possibly these literatures to be the milestones. Over the last decade, epigenetic research are mainly in three aspects: (1) DNA mehtylation--C11, C52, C59, C60 and C63 (2) histone modifications --C25 (3) noncoding RNA(ncRNA) --C40,C41.

In Table 4, we listed the top 40 significant articles, which are first sorted by clusters, then sorted in chronological.

Table 4. Top 40 milestone references of epigenetic between 2002 and 2013

No.	*frequ-ency*	*burst*	*centrality*	*geomean*	*year*	*References*
1	623	16.16	0.18	1.71	**1992**	**LI E，1992，Targeted mutation of the DNA methyltransferase gene results in embryonic lethality，CELL**，V69，P915
2	482	8.05	0.14	1.06	1992	**Frommer M, McDonald LE, et al. A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands .Proc, Natl.Acad.Sci.USA.1992 80:1579-83.**
3	351	23.34	0.07	1.28	**1993**	**LI E，1993，Role for DNA methylation in genomic imprinting，NATURE，V366，P362**
4	1258	14.19	0.11	1.25	**1996**	**Herman JG，1996，Methylation-specific PCR: A novel PCR assay for methylation status of CpG islands，** P NATL ACAD SCI USA，V93，P9821
5	521	39.87	0.15	2.45	**1998**	**Nan XS，1998，Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex，NATURE，V393，**

6	333	28.8	0.05	1.20	**1998**	**Herman JG，1998，Incidence and functional consequences of hMLH1 promoter hypermethylation in colorectal carcinoma，P NATL ACAD SCI USA，V95，P6870**
7	449	68	0.01	0.82	**1998**	**Baylin SB，1998，Alterations in DNA methylation: a fundamental aspect of neoplasia，ADV CANCER RES，V72，P141**
8	691	66.87	0.1	2.59	**1999**	***Jones PA，1999，Cancer epigenetics comes of age，NAT GENET，V21，P163**
9	432	32.42	0.12	1.97	**1999**	***Bird AP，1999，Methylation-induced repression - Belts, braces, and chromatin，CELL，V99，P451**
10	834	3.91	0.15	0.77	**1999**	**Okano M，1999，DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development，CELL，V99，P247**
11	421	48.01	0.03	1.20	**2000**	**Dammann R，2000，Epigenetic inactivation of a RAS association domain family protein from the lung tumour suppressor locus 3p21.3，NAT GENET，V25，P315**
12	452	48	0.03	1.20	**2000**	***Baylin SB，2000，DNA hypermethylation in tumorigenesis，TRENDS GENET，V16，P168**
13	266	19.04	0.07	1.15	**2000**	**Rountree MR，2000，DNMT1 binds HDAC2 and a new co-repressor, DMAP1, to form a complex at replication foci，NAT GENET，V25，P269**
14	286	16.26	0.05	0.90	**2000**	**Fuks F，2000，DNA methyltransferase Dnmt1 associates with histone deacetylase activity，NAT GENET，V24，P88**
15	435	13.74	0.1	1.17	**2000**	**Costello JF，2000，Aberrant CpG-island methylation has non-random and tumour-type-specific patterns，NAT GENET，V24，P132**
16	908		0.23	0.23	**2000**	***Strahl BD，2000，The language of covalent histone modifications，NATURE，V403，P41(progress)**
17	327	46.98	0.07	1.81	**2001**	**Tamaru H，2001，A histone H3 methyltransferase controls DNA methylation in Neurospora crassa，NATURE，V414，P277**
18	481	24.76	0.09	1.49	**2001**	**Bannister AJ，2001，Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain，NATURE，V410，P120**
19	1350		0.09	0.09	**2001**	**Jenuwein T，2001，Translating the histone code，SCIENCE，V293，P1074**
20	595	23.62	0.04	0.97	**2001**	**Esteller M，2001，A gene hypermethylation profile of human cancer，CANCER RES，V61，P3225**
21	334	22.49	0.18	2.01	**2002**	**Jackson JP，2002，Control of CpNpG DNA methylation by the KRYPTONITE histone H3 methyltransferase，NATURE**，V416，P556
22	407		0.23	0.23	**2002**	**Cao R，2002，Role of histone H3 lysine 27 methylation in polycomb-group silencing，SCIENCE，V298，P1039**
23	1033		0.14	0.14	**2002**	***Bird A，2002，DNA methylation patterns and epigenetic memory，GENE DEV，V16，P6**
24	446	13.47	0.09	1.10	**2002**	***Li E，2002，Chromatin modification and epigenetic reprogramming in mammalian development，NAT REV GENET，V3，P662**
25	1220	20.34	0.08	1.28	**2002**	***Jones PA，2002，The fundamental role of epigenetic events in cancer，NAT REV GENET，V3，P415**
26	478	17.14	0.02	0.59	**2004**	***Feinberg AP，2004，The history of cancer epigenetics，NAT REV CANCER，V4，P143**
27	355	12.87	0.05	0.80	**2005**	**Weber M，2005，Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells，NAT GENET，V37，P853**
28	338	3.95	0.04	0.40	**2005**	**Fraga MF，2005，Loss of acetylation at Lys16 and trimethylation at Lys20 of histone H4 is a common hallmark of human cancer，NAT GENET，V37，P391**

29	678	34.59	0.08	1.66	**2006**	**Bernstein BE, 2006, A bivalent chromatin structure marks key developmental genes in embryonic stem cells, CELL, V125, P315**
30	970	118.27	0.03	1.88	**2007**	***Kouzarides T, 2007, Chromatin modifications and their function, V128, P693**
31	319	36.58	0.08	1.71	**2007**	**Weber M, 2007, Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome, NAT GENET, V39, P457**
32	551	71.41	0.01	0.85	**2007**	***Barski A, 2007, High-resolution profiling of histone methylations in the human genome, CELL, V129, P823(resource)**
33	631	68.12	0.01	0.83	**2007**	***Jones PA, 2007, The epigenomics of cancer, CELL, V128, P683**
34	467	57.61	0.01	0.76	**2007**	**Mikkelsen TS, 2007, Genome-wide maps of chromatin state in pluripotent and lineage-committed cells, NATURE, V448, P553**
35	310	73.27	0.07	2.26	**2008**	**Meissner A, 2008, Genome-scale DNA methylation maps of pluripotent and differentiated cells, NATURE, V454, P766**
36	205	38.83	0.04	1.25	**2008**	**Cokus SJ, 2008, Shotgun bisulphite sequencing of the Arabidopsis genome reveals DNA methylation patterning, NATURE, V452**, P215
37	405	90.61	0	0.87	**2008**	***Esteller M, 2008, Molecular origins of cancer: Epigenetics in cancer, NEW ENGL J MED, V358, P1148**
38	309	116.12	0.05	2.41	**2009**	**Lister R, 2009, Human DNA methylomes at base resolution show widespread epigenomic differences, NATURE, V462, P315**
39	202	70.66	0.08	2.38	**2009**	**Tahiliani M, 2009, Conversion of 5-Methylcytosine to 5-Hydroxymethylcytosine in Mammalian DNA by MLL Partner TET1, SCIENCE, V324, P930**
40	145	50.49	0.9	6.74	**2009**	**Kriaucionis S, 2009, The Nuclear DNA Base 5-Hydroxymethylcytosine Is Present in Purkinje Neurons and the Brain, SCIENCE, V324, P929**

Note : geomean refers to the geometric mean of the burst and the centrality.

Conclusions

By mapping the intellectual structure of epigenetic research over the last decade, we found epigenetics has developed in a new trend of interdisciplinary study and cross-disciplinary development in recent years and identified DNA methylation and histone modification are mainly the epigenetic research hotspots over the past 10 years; ncRNA may likely be one main research front over the last 5 years. Meanwhile, we detected top 40 significant literatures over the last decade.

Analyzing the hotspots, fronts and the top40 literatures, we conclude epigenetics research has definitely developed over time, possibly these literatures to be the milestones. For the methods of the citation frequency, burst and structure centrality, our results suggest important articles may not be judged only by one approach, but considering the geometric mean of the burst and betweenness centrality value as a comprehensive evaluation of the influence of the literatures may likely be more reasonable.

To be supplement statement, for the vast amounts of our original database, using only top40 as a summary is inevitably a bit incomprehensive. However, visual analytics of the literature provides a valuable, timely, repeatable and flexible approach so as to track the development process of a scientific field and identify significant theory and technology, and it could be useful for science researchers and clinicians who are not quite familiar with the domain to choose correct direction and identify entry points to have a start.

References

Baylin SB.(1998). Alterations in DNA methylation: a fundamental aspect of neoplasia. *Adv Cancer Res 72*, 141-146.

Baylin SB, Herman JG. (2000). DNA hypermethylation in tumorigenesis.*Trends Genet* 16, 168-174.

Bird A.(2002).DNA methylation patterns and epigenetic memory, DNA methylation patterns and epigenetic memory. *Gene Dev* 16, 6-21.

Ball MP, Li JB , Gao Y, et al., 2009. Targeted and genome-scale strategies reveal gene-body methylation signatures in human cells. Nature Biotechnology 27, 361-368.

Borner K, Huang WX, Linnemeier M, et al.(2010). Rete-netzwerk-red: analyzing and visualizing scholarly networks using the Network Workbench Tool. *Scientometrics* 83,863-876.

Chen CM.(2004). Searching for intellectual turning points: progressive knowledge domain visualization. *Proc Natl Acad Sci USA* 101,5303-5310.

Chen CM.(2006). CiteSpace II: detecting and visualizing emerging trends and transient patterns in scientific literature. *J Am Soc Inf Sci Technol* 57,359-377.

Chen C, Song IY, Yuan XJ, Zhang J.(2008). The thematic and citation landscape of data and knowledge engineering (1985 -- 2007). *Data Knowledge Eng* 67,234-259.

Chen CM, Ibekwe-SanJuan F, Hou JH.(2010). The structure and dynamics of cocitation clusters: a multiple-perspective cocitation analysis. *J Am Soc Inf Sci Technol* 61,1386-1409.

Chen, C.(2011). Turning Points: The Nature of Creativity, *Springer and Higher Education Press* 5,127-129.

Cobo MJ, Lopez-Herrera AG, Herrera-Viedma E, Herrera F.et al.(2011). Science mapping software tools: review, analysis, and cooperative study among tools. *J Am Soc Inf Sci Technol* 62 (7): 1382-402.

Chen Chaomei, Hu Zhigang, Liu Shengbo and Tseng Hung.(2012). Emerging trends in regenerative medicine: a scientometric analysis in CiteSpace. *Expert Opinion on Biological Therapy* 12,593-608.

Chen C.(2012). Predictive effects of structural variation on citation counts. *J Am Soc Info Sci Technol* 63,431-449.

Chen C.(2012). Predictive effects of structural variation on citation counts. *J Am Soc Info Sci Technol* 63,431-449.

Garfield E.(2004). Historiographic mapping of knowledge domains literature. *J Inf Sci 30(2)*, 119-145.

Goldberg AD, Allis CD, Bernstein E,et al., 2007. Epigenetics: a landscape takes shape. Cell 128(4), 635-638.

Herman JG, Graff JR, Myohanen S.(1996). Methylation-specific PCR: A novel PCR assay for methylation status of CpG islands. *P Natl Acad Sci USA* 93, 9821-9826.

Hammond SM, Bernstein E, Beach D, et al.(2000). An RNA-directed nuclease mediates post-transcriptional gene silencing in Drosophila cells. *Nature* 6775, 293-296.

Jenuwein T, Allis CD.(2001). Translating the histone code. *Science* 293, 1074-1080.

Jones PA, Laird PW, 1999. Cancer epigenetics comes of age. Nat Genet 21, 163-167.

Jones PA, Baylin SB.(2002).The fundamental role of epigenetic events in cancer. *Nat Rev Genet* 3, 415-428.

Jimenez-Marin A, Collado-Romero M, Ramirez-Boo M, et al., 2009. Biological pathway analysis by Array Unlock and ingenuity pathway analysis. BMC Proc 3(Suppl 4),S6.

Kouzarides T.(2007).Chromatin modifications and their function. *Cell* 128， 693-705.

Kriaucionis S， Heintz N.(2009). The Nuclear DNA Base 5-Hydroxymethylcytosine Is Present in Purkinje Neurons and the Brain. *Science* 324, 929-930.

Leydesdorff L, Schank T.(2008). Dynamic animations of journal maps: indicators of structural changes and interdisciplinary developments. *J Am Soc Inf Sci Technol* 59,1810-1818.

Lister R， Pelizzola M , Dowen RH, et al.(2009). Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature* 462， 315-322.

Morris SA, Yen G, Wu Z, Asnake B, et al.(2003). Timeline visualization of research fronts. *J Am Soc Inf Sci Technol* 55,413-422.

Nan XS， Ng HH, Johnson CA,et al.(1998). Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. *Nature* 393， 386-389.

Shannon P, Markiel A, Ozier O, et al., 2003.Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 13(11),2498-2504.

Small H.(1973).Co-citation in the scientific literature: a new measure of the relationship between two documents. *J Am Soc Info Sci* 24,265-269.

Strahl BD， Allis CD.(2000).The language of covalent histone modifications.*Nature* 403,41-45.

Tahiliani M， Koh KP, Shen YH,et al., 2009. Conversion of 5-Methylcytosine to 5-Hydroxymethylcytosine in Mammalian DNA by MLL Partner TET1.Science 324,930-935.

van Eck NJ, Waltman L.(2010). Software survey: VOSviewer, a computer programfor bibliometric mapping. *Scientometrics* 84,523-538.

Zeyuan Liu, Yue Chen, Haiyan Chen.(2008). Mapping Knowledge Domains Methods and Application. 1st ed. *People Publishing House*. *Peking*, 34-35, 68-81.

Acknowledgments

The research reported in this paper is supported by the soft science research projects in Shanxi Province under grant numbers 2013041082-02 and 2012041068-01 and by National Natural Science Foundation of China under grant number 71103114.

Characteristics of the development of NB converging technology

Hou Haiyan*, Zhao Nannan*, Zhang Shanshan*, Liang Yongxia** and Hu Zhigang*

*WISE Lab, Dalian University of Technology, Dalian, China 116085
huzhigang.cn@gmail.com

**Editorial department, Journal of China Science and Technology, Beijing 100190

Abstract

In the 21st century, the development of converging technology attracts more and more attention gradually. Firstly, we give a brief introduction about the-state-of-the-art of nano-bio convergence technology (NB). Then we dig out the key technology of NB converging technology, namely farmdoc-processes, apparatus-nanotechnology (general) (B11-C12). Then we use the journal and conference articles proportion method, Fisher-pry model, Gong lang hereby (GomPertz) curve model, Logical growth curve model to to study the characteristic of the development of NB converging technology. We notice that NB converging technology is in a period of rapid development, and has not yet reached the level of maturity. So we choose Fisher-pry model to preliminary predict that NB converging technology might reach a level of maturity in 2019.

Keywords: NB converging technology; farmdoc-processes, apparatus-nanotechnology (general) (B11-C12); journals and conference articles proportion method; Fisher-pry model; Gong lang hereby (GomPertz) curve model; Logical growth curve model

Introduction

The concept of “converging technology” was initially proposed in a conference sponsored by the U.S. department of commerce technology administration, the national science foundation (NSF) and the national science and technology commission of Nano science and engineering and technology board (NSTC-NSET) in which scientists, government officials and industrials all participated. In this conference, converging technology refers to the coordination and integration of the big four field of science and technology including Nano science and technology, biotechnology and biological medicine (including genetic engineering), information technology (including advanced computer and communication), cognitive science (including cognitive neuroscience) which are rapidly developing in 21st century. Its simplified English coupling is "NANO - BIO - INFO - COGN", abbreviation for "NBIC"[1]. Converging technology of Nano science and technology and biotechnology are abbreviated to “NB”.

Many experts believe that we human being will entry a century of converging technology in the 21st century, and biotechnology will be the core technology which will become a new economy growth point after the information technology [2]. With the progress of science, biotechnology develops to high-tech and gradually entry the stage of nano-biotechnology. The article briefly introduces the research status and development prospect of converging technology, and emphasizes a further study of characteristics of the development of NB converging technology.

1. Data and methods

Firstly, we download the data by using the search strategy of TS=(("selfassembl*" OR "self assembl*" OR "atom force microscop*" OR atom-force-microscope* OR "scanning tunneling microscop*" OR "scanning-tunneling-microscop*" OR "atomistic simulation" OR "molecular device" OR "molecular electronics" OR "molecular modeling" OR "molecular motor" OR "molecular sensor" OR"molecular simulation" OR "quantum computing" OR "quantum dot*" OR "quantum effect*" OR nano*) NOT (nano3* OR nano2 OR nanosomia* OR nanook* OR nanosaurus*)) (research strategy from nano-science keywords list of NSF for NSE award statistics) in Web of Science and Derwent Innovation Index respectively. Then, we refine the data belonging to biotechnology subject category. Finally, we get 32932 articles from Web of Science and 7443 patents from Derwent Innovation Index.

Using these data, we make a series of Excel charts to show the state-of-arts of the NB converging technology. Then, based on social network analysis and co-occurence analysis, we dig out the key technology of NB converging technology. Finally, we do a fit with Gong lang hereby (GomPertz) curve mode and Logical growth curve model respectively.

It's much easier to retrieve the data of journal articles and conference papers. Using the same search strategy in web of science and setting appropriate options, then refine the data belonging to biotechnology subject category. After that, if we define journal articles, we could get the number of journal articles, otherwise, if we define conference articles (proceeding paper、meeting abstract、meeting summary), we could get the number of conference articles. We apply those above data in the journals and conference articles proportion method to investigate the characteristics of the development of NB converging technology.

The focus of SCI and EI is different. SCI places emphasis on basic research; however EI focuses on applied research. We use the same research strategy and set appropriate options, then define with "bio*". Finally, we get the number of EI articles per year. The sum of SCI articles' number and EI articles' number per year is the data source of Fisher-pry model.

2. The-state-of-the-art of NB converging technology

Study on Nano materials is the hottest and the most promising research field in the field of materials science. Nano technology is applied in many fields including ceramics, microelectronics, bioengineering, photoelectricity, medicine. Nano-biotechnology is the frontier and hot issue in the field of biotechnology. It has been widely used and has a well-defined industrial prospect in medical and health area, especially Nano drug carrier, Nano-biosensors and imaging technology as well as micro intelligent medical apparatus and instruments which will play an important role in disease diagnosis, treatment and health care.

At present, nano-biotechnology on the international has achieved a certain degree of development in the field of medicine. America, Japan, Germany and other countries have made nano-biotechnology as the research priorities and focused on its development. Compared with these advanced countries, China started late on nano-biotechnology. But during the "ninth five-year", "863 plan" launched the national Nanotechnology revitalization plan and during "tenth five-year"in "863 plan" nano-biotechnology was listed as priority support development project. The research of nano-biotechnology mainly focuses on the following directions, including nano-biomaterials, nano-bio devices and the application of nano-bio convergence technology in clinical diagnosis[3]. Based on the data, we make a series of excel charts to explain the state-of-the-art of converging technology visually.

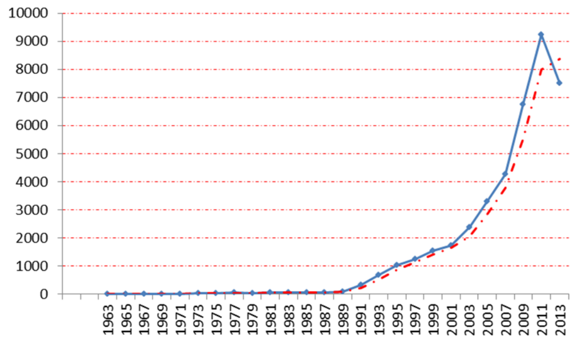


Figure 1 Growth figure of the number of scientific papers of NB converging technology

As shown in figure 1, the number of scientific papers of converging technology increased from 1963 to 2013. There was only two articles in 1963 and until 1989 the number was below 100. However, the number suddenly increased from 1991 and it reached 1039 in1995. Since then, the number showed a trend of exponential growth and lasted until now. This trend reveals that people pay more and more attention to the development of NB converging technology.

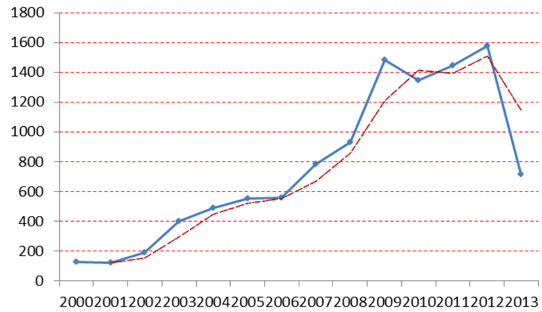


Figure 2 Change trend of patent applications of NB converging technology

We analysis the patent application data after 2000. As shown in figure 2, patent application increased slowly from 2000 to 2002, there was 126 in 2000 and 187 in 2002. From 2003, the number became exponential growth and there was 400 in 2003 that was 3.17 times to the number in 2000. It reached a small peak in 2009 and was 1482. But patent application decreased in 2010, then it increased. This reflects that NB converging technology is being integrated continuously.

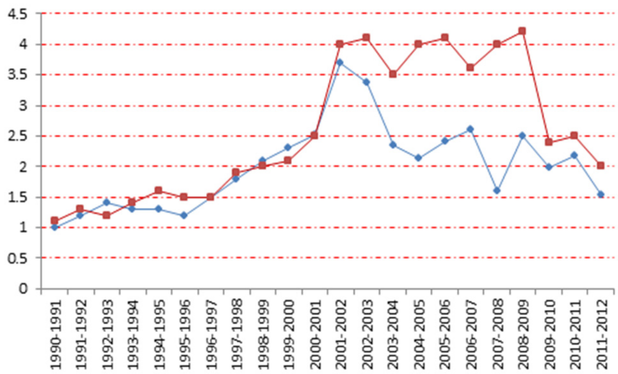


Figure 3 Change trend of growth rate of the number of scientific papers (red line) and patent application of NB converging technology (blue line)

$$a = [1/(t_2 - t_1)]\sum_{t=t_1}^{t_2}[Y(t+1)/Y(t)] \qquad (1)$$

Formula (1) is used to calculate growth rate of the number of scientific papers and patent application of NB converging technology. Here Y(t) indicates the number of scientific papers or patent application of NB converging technology in the "t"th year. In figure 3, the growth rate of the two types of papers have same trend which reflects that theoretical research are synchronized with practical application and the development is coordination. This fully shows that people pay attention to the development of NB converging technology which is in coordinated and healthy period. From 2000 to 2012, the growth rate is falling down. During 2001 to 2002, the growth rate is the largest. This reflects the growth spread reaches a peak which is related with proposition of the concept of converging technology in 2001. This concept is initially proposed in a conference sponsored by the U.S. department of commerce technology administration, the national science foundation (NSF) and the national science and technology commission of Nano science and engineering and technology board (NSTC-NSET) in which scientists, government officials and industrials all participated.

3. Characteristics of the development of NB converging technology

3.1 Dig out the key technology of NB converging technology

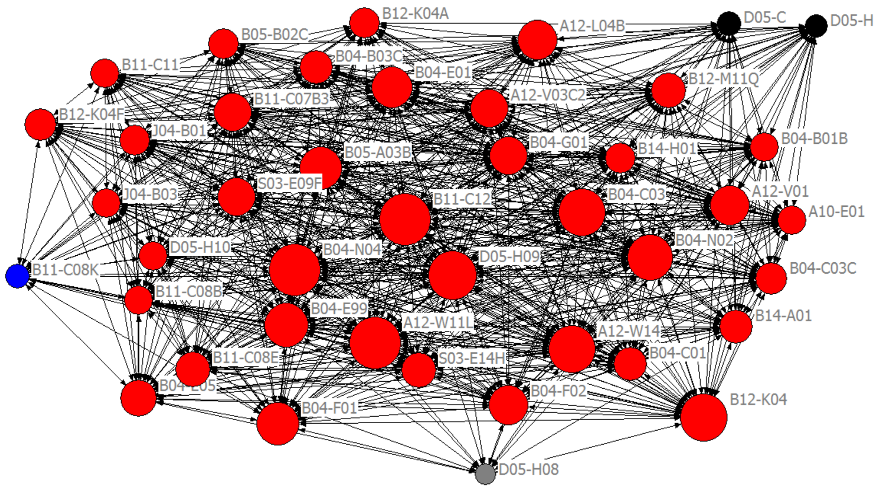


Figure 4 Main distribution area of NB converging technology

Use bibexcel to process 7443 patent data, extract derwent manual code and then use ucinet to do co-occurence analysis. As shown in the figure 4, NB converging technology are divided into three categories (here we denote using MC codes). One is mainly concentrated on B11-C12 (nanotechnology[general]), D05-H09 (testing and detection [exc. bacteria, fungi, viruses]), B12-K04 (diagnosis and testing), B04-C03 (polymers[general]), A12-W11L ((immobilised)enzymes or microorganisms, microbiology (polymer use)) and B12-M11Q (Nano formulation). The second one is B11-C08K(other analytical apparatus where the apparatus is the novelty of the invention). The third one is D05-C (Chemicals by fermentation (biosynthesis) [others;general]), D05-H (microbiology, laboratory procedures [general and others]). Farmdoc — processes or apparatus — Nanotechnology (general) (B11-C12) is in the core position and it is connected closely with other technologies so that it is bridge to link others. As shown in table 1, its patent quantity is the largest which reaches 141 and the degree is 41, and betweenness centrality is 6.445 that is equal to protein or polypeptide of undefined origin and polymer application technology of microorganism or enzyme. To sum up, no matter from any perspective, agricultural processes or facilities Nanotechnology is the first position. So it is the key technology.

Table 1 Areas of NB converging technology and its patent quantity, degree, and betweenness centrality

MC Code	Domain	Patent Quantity	Degree	Betweenness centrality
B11-C12	Nanotechnology [general]	141	41	6.446
D05-H09	testing and detection [exc. bacteria, fungi, viruses]	126	40	5.681
B12-K04	diagnosis and testing	77	39	5.57
B04-E01	nucleic acid general and other	68	39	4.166
B04-N04	protein / polypeptide of undefined origin (no sequence) (1994 —)	58	41	6.446
B04-C03	polymers[general]	51	40	5.386
A12-V01	medcines,pharmaceuticals	51	37	3.764

A12-W11L	(immobilised)enzymes or microorganism, microbiology (polymer use) (1986 －)	48	41	6.446
A12-W14	Nanotechnology	46	40	5.386
B04-G01	antibody defined in terms of antigen general and other (1994 －)	46	39	3.568
B12-M11Q	Nanoformulations	46	36	3.468
S03-E09F	immunoassay techniques and biological indicators (2005 －)	45	38	3.546
D05-H	microbiology, laboratory procedures [general and others]	43	40	5.681
D05-H10	fixing biological substances or cells to a carrier and the carriers themselves	39	36	3.468
B04-N02	animal protein/polypeptide (no sequence)	35	36	5.216

3.2 Technology maturity

3.2.1 Journals and conference papers proportion method[4]

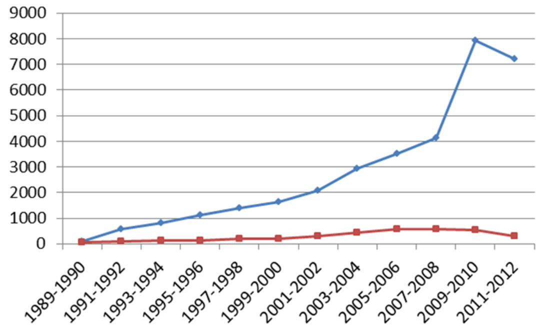


Figure 5 Change trend of the number of journal papers (blue line) and conference papers (red line)

Roper assumes that ratio of the number of journal papers to conference papers can determine technology maturity. When the number of conference papers is more than journal papers, it indicates the new (emerging) technology is still in the debate and the technology is far from maturity. When the number of journal papers is much more than conference papers, it indicates the technology is closing to maturity gradually. As shown in figure 5, there is no first condition, because NB converging technology are proposed late. And conference papers is not much more so that the number of conference papers is initially less than journal papers. This is a special case, so we can determine whether it is mature according to the ratio of the number of conference papers and journal papers to the total papers. The data source is web of science and their values are shown in table 2. Figure 6 shows their ratio.

Table 2 The number of scientific papers and patent application of NB converging technology

Time Partition	Journal Papers	Conference Papers	Sum
1989-1990	107	45	152
1991-1992	578	81	658
1993-1994	810	126	936
1995-1996	1116	129	1245
1997-1998	1400	185	1585
1999-2000	1629	183	1812
2001-2002	2093	303	2396
2003-2004	2923	451	3374
2005-2006	3499	564	4063

2007-2008	4122	574	4696
2009-2010	7928	524	8452
2011-2012	7207	312	7519

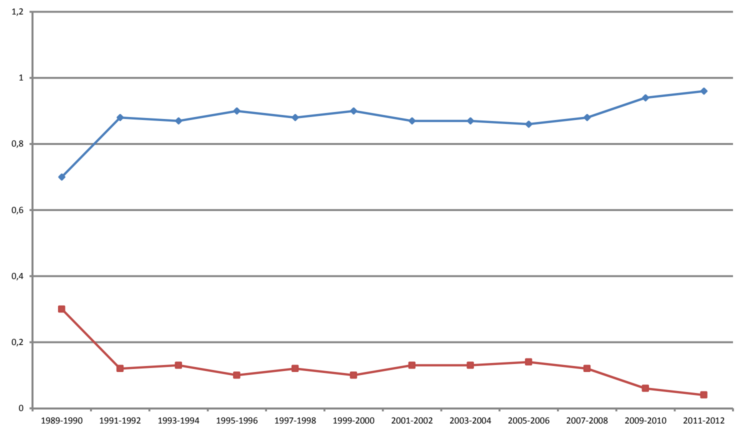


Figure 6 Change trend of ratio of journal papers (blue line) and conference papers (red line)

As shown in figure 6, during 1989 to 1990, the number of journal papers is equal to conference papers which indicates that this technology is initially proposed and is still in debate. After 1991, the number of journal papers is much more than conference papers which indicates that more and more people begin to study the technology and study further even include wider aspects, but it does not reach maturity. Only when the ration of journal papers is 1 and the ratio of conference papers is 0, can it indicate the technology reach maturity. Figure 6 shows the journal curve is closing to 1 and conference curve is closing to 0 which indicates that NB converging technology have much room to develop.

3.2.2 Comparing among the three models

3.2.2.1 Fisher-pry model[4]

To judge whether a technology is mature, analysis based on articles' quantity has some deficiencies. The data in the analysis needs to be standardized to meet the S curve of the development change of technology. Fisher-pry model is a ideal method to judge whether a technology is mature. Fisher and Pry published a article about a model of depicting technology change in 1971. The model is easy, but it is effective to judge whether a new technology can replace an old technology which is competitive with it and also can estimate the ratio of new technology replace old technology. This model with a little modification and appropriate data can define the degree of technology improvement and estimate whether a technology is mature.

Table 3 Articles of NB converging technology in each year (the data in 2013 is predicted according to the model)

Year	SCI	EI	Sum
1990	69	1	70
1991	247	3	250
1992	314	2	316
1993	315	9	324
1994	484	15	499
1995	511	12	523

1996	529	10	539
1997	697	12	709
1998	717	12	729
1999	788	25	813
2000	814	40	854
2001	906	38	944
2002	1114	79	1193
2003	1240	146	1386
2004	1627	261	1888
2005	1744	459	2203
2006	1853	687	2540
2007	2028	1081	3409
2008	2171	1724	3895
2009	3673	2250	5923
2010	4175	2388	6563
2011	3606	2891	6497
2012	3601	2942	6543
2013			6370
2014			6650
2015			6580
2016			6650
2017			6720
2018			6790
2019			6930
2020			6930
2021			7000

The following is the analysis of the maturity of NB converging technology by Fisher-pry model. To map the Fisher-pry model, firstly, standardize the data in table 3 by equation (1).

$$f = \frac{1}{1 + c * e^{-b*t}} \quad (1)$$

In formula (1), b and c is constant, b decides the shape of the curve, c decides the position of the curve. f is substitution rate, $0<f<1$, which can get by formula (2).

$$f = \frac{Y}{L} \quad (2)$$

In formula (2), Y is the number of published papers in a certain year, L is the largest number of published papers to be estimate. In this case , according to the investigate and forecasting, L=7000.

$$\ln \frac{f}{1-f} = b * t - \ln C \quad (3)$$

Based on linear regression method (this case uses SPSS software), we can calculate b and c, b=0.3, $c=e^{601.5}$. Therefore, we can get figure 7 by formula (1).

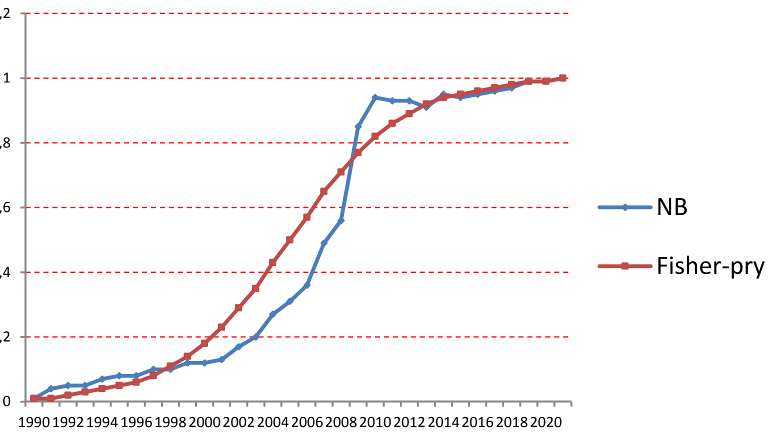


Figure 7 Fisher-pry model

As shown in figure 7, the curve is not close to 1 in 2013 which reflects that NB converging technology is still in rapid development and far from maturity. The curve is close to 1in2019, so from the perspective of bibliometrics we can predict that it would reach maturity in 2019. There needs further argument with experts in this field on its accuracy, but not explained here.

3.2.2.2 Gong lang hereby (GomPertz) curve model[5]

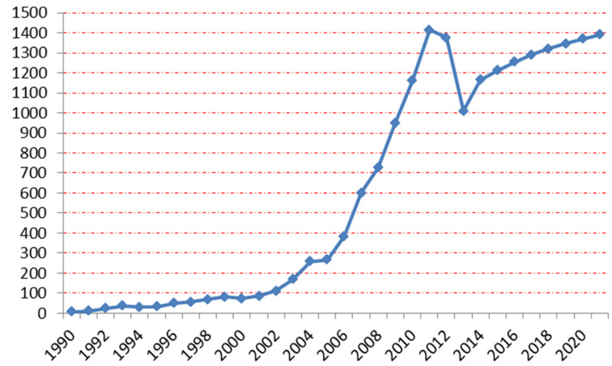


Figure 8 Gong lang hereby (GomPertz) curve model

Gong lang hereby curve model is:

$$f = Fe^{-ae^{-bt}} \qquad (4)$$

In formula (4): f indicates predictive variable which increases over time; F is the upper limit of f. In this case, F=1500 according to relevant investigate and prediction. a and b are the coefficients calculated by the “fitting” between curve and data. e is the natural logarithm of a bottom.

By the formula (4), get

$$\ln[-\ln\frac{f}{F}] = -bt + \ln a \qquad (5)$$

Based on linear regression method (this case uses SPSS software), we can calculate a and b, $a=e^{281}$, b=0.14. Therefore, we can get figure 8 by formula (4). As shown in figure 8, Gong lang hereby curve model is similar to Fisher-pry model. According to the character of the model, the inflection point is in the position that $t = \ln\frac{a}{b}$ =2007, that we can distinguish maturity of technology from a macroscopic view according to the speed of development: technology which is before the inflection point is in its infancy or growth period; technology which is after the inflection point is in maturity period or out of phase. In 2013, the curve is still in growth period, so NB converging technology is rapid developing and does not reach maturity. The curve has leveled off in 2020 and we can predict that the technology would reach maturity in 2020.

3.2.2.3 Logical growth curve model[5]

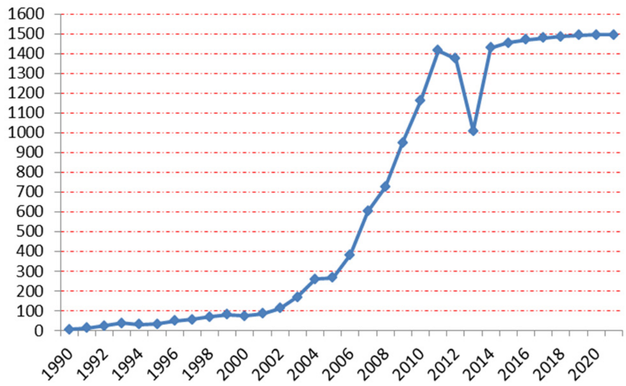


Figure 9 Logical growth curve

Logical growth curve model is:

$$y = \frac{k}{1 + ae^{-bt}} \tag{6}$$

In formula (6), y is predictive variable (function), k is absolute rating, a and b are model parameter, t is time variable, k is the limit value. In this case, according to relevant investigate and prediction, k=1500. If $t -> -\infty$, then $y -> 0$; if $t -> +\infty$, then $y -> k$.

Get by formula (6),

$$\ln\frac{k - y}{y} = \ln a - bt \tag{7}$$

Based on linear regression method (this case uses SPSS software), we can calculate a and b, a=e^{863}, b=0.43. Therefore, we can get figure 9 by formula (6). According to the character of the model, the inflection point is in the position that $t = \ln\frac{a}{b}$ =2007, that we can distinguish maturity of technology from a macroscopic view according to the speed of development: technology which is before 2007 is in its infancy or growth period; technology which is after 2007 is in maturity period or out of phase. As shown in the figure, in 2013, NB converging technology is rapid developing. The curve has leveled off in 2019, so we can predict that the technology would reach maturity in 2019.

3.2.2.4 Choice of model—Minimum residual sum of squares identification method[5]

Minimum residual sum of squares identification method takes the curve of optimal as principle. Compute residual sum of squares of various of curves by time series. Choose the curve with the minimum residual sum of squares as predictive model.

Suppose: $y_1, y_2 \cdots\cdots y_n$ is historical observations, $\bar{y}_1, \bar{y}_2 ......, \bar{y}_n$ is predictive value got by the model. The difference between actual value and predictive value is

$$e_i = y_i - \bar{y}_i (i = 1, 2, 3.....n) \quad (8)$$

For the same set of data points, different curve has different residual sum of squares. Therefore choosing the curve whose Q value is minimum as predictive model, then we can get minimum fitting error.

After calculating, the residual sum of squares of Fisher-pry model is Q1=11577, of Gong lang hereby model is Q2=843292.8, of Logical growth curve model is Q3=335788.5, we get,

$$Q1 < Q3 < Q2$$

Therefore, Fisher-pry model is the most accurate among the three models. Finally, we choose Fisher-pry model to forecast the maturity of NB converging technology.

4. Conclusion

From the perspective of technology maturity, NB converging technology does not reach maturity and have much room to develop. Studying and manipulating biomacromolecule in Nano level is a challenge to human recognition ability and research ability. Life movement as the highest form of movement in nature contains a lot of secrets. There are many so far which have not been known. Fifty years ago, Austrian physicist proposed in a book whose name is "What Is Life": microworld of life movement is molecular machine. Molecular biology research at present fully proved physicists' prediction. Observe and study in nanometer level, biomacromolecule which takes DNA molecule as core is indeed self-assembly molecule machine. Recently scientists have conducted test: linking one end of two complementary DNA single chains to nanometer gold parties, DNA chains complementary combining after mixture, nanometer gold parties will realize self-assembly. Although nanobiotechnology is just emerging, it has enormous potential to develop to promote the construction of human material civilization as a new emerging technology.

Nanotechnology field is one of the most promising fields in 21st century. Nano-biotechnology is the frontier and hot issue in national biotechnology field, and have wide application and specific industrial prospect in medicine and health care field. Currently research fields of nano-biotechnology mainly focus on the following directions, including nano-biomaterials, nano-bio devices and the application of nano-bio convergence technology in clinical diagnosis. Through research for NB converging technology, we find NB converging technology's main domain is Nanotechnology (general) (B11-C12), testing and detection [exc. bacteria, fungi, viruses](D05-H09), diagnosis and testing(B12-K04), natural products (polymers[general](B04-C03), (immobilised)enzymes or microorganisms and microbiology (polymer use)(A12-W11L) and Nano formulation(B12-M11Q). The key domain is farmdoc — processes or apparatus — Nanotechnology (general) (B11-C12). Through research of maturity of NB converging technology, we find it is far from maturity and have much room to develop. We can predict it may reach maturity in 2019 by the Fisher-pry model. In future, NB converging technology will do further study for biomacromolecule which take DNA molecule as core. This will promote human beings' development. Human will entry a biology century in 21st century, and biotechnology will be core technology which will be a new economy growth point after information technology. With the progress of science, biotechnology develops to high-tech and gradually entry the stage of nano-biotechnology[6].

Acknowledgement

The article is a periodical research result of the major project of the national social science fund (No. 12&ZD117, 14BTQ030, 10CZX011), the project of the natural science fund (71373253, 71103022), the project of the innovation approach (2013IM010100), the project of Chinese degree and graduate education society (2013Y12), the reform project of graduate education of Dalian University of Technology (JG2013002), the planning project of social science of Liaoning province (L13BGL014).

Corresponding author: Zhigang Hu, huzhigang.cn@gmail.com.

Reference

[1] Zhang Ning, Luo Changkun. "Converging Technology" and Its Impact to S&T Management[J]. *R&D Management*, 2005, 17(5): 97-100. (In Chinese)

[2] Zhang Yangde. The prospects of Nano-biotechnology[J]. *China Chemicals*, 2003, 10:004. (In Chinese)

[3] Zhu Donghua, Yuan Junpeng. The studies of technology monitoring indicators[J]. Studies In *Science of Science*, 2003, 21(4): 419-422. (In Chinese)

[4] Lou Yan, Fu Xiaoyang, Huang Lucheng. A study on technology maturity based on bibliometrics[J]. *Statistics and Decision*, 2010, 19:032. (In Chinese)

[5] Zhang Huangao, Zhao Wenyan, Tan Runhua. Technique of Product Technology Maturity Mapping Based on Patent Analysis and Software Design[J]. *China Mechanical Engineering*, 2006, 17(8): 823-827. (In Chinese)

[6] Jin Gang, Ying Peiqing. Nano-biotechnology[J]. *Nature Magazine*, 2001(4): 211-214+184-248. (In Chinese)

Data Fetching and Group Characteristics Analysis Based on Sina Microblog

Jiang Chunlin, Liu Xue and Zhang Liwei

WISE Lab, School of Public Administration and Law, Dalian University of Technology, Dalian 116085 China
chunlinj7873@163.com, 314014155@qq.com, taiyiliwei@126.com

Abstract

With the development of Web 2.0, SNS becomes the most popular and important channel of publishing, communicating and spreading information, most of the internet users have built the habits of surfing on the SNS websites every day. Now in China, Sina Microblog is the SNS platform which has most of the users in all of SNS. In this article, the API port provided by Sina and a self-designed web spider are used as the data extraction method to get the user related information. In addition, in order to investigate the network characteristics of the Sina Microblog users. A deep data mining is conducted by the knowledge of Webometrics and complex network. The result of this research may lead some directions to explore the business value of the SNS.

Keywords: SNS; Webometrics; Complex Network; Group Characteristic; Sina Microblog

Introduction

The SNS (Social Network Site) is a world-wide popular tool, and has become an extreme important information exchanging platform. Compared with the traditional media, SNS has the characteristics of wide range resources, fast spreading speed, huge influence scope and timeliness. A growing number of social events are published on SNS at the first time, and soon become the center of the attention. The marketing events and the public relations affairs in the SNS are very essential to the companies. So we can definitely say that SNS not only bringing the revolution to the information and technology field, but also changing the humans' life style, interaction way and the thinking way. Most importantly, it has a tremendous impact to each kind of aspect of people's life and the development of human being.

As a new network application form in Web 2.0 era, microblog has gotten a rapid development. Users update their recent information in about 140 words, and spread to their fans quickly, Paving a way not only meet the rapid pace of life, but also convenient in sending and sharing the information (Westman S & Freund L, 2010). According to the official statistics of Sina Microblog, until the end of March 2013, Sina Microblog has 536 million registered users, 60.2 million users use Sina Microblog every day. The data reflects that Sina Microblog can represent the chief social networking platform of China.

In the recent years, the research of data analysis based on SNS has got a widely attention. Ye Wu and Jurgen Kurths (Ye Wu & Jurgen Kurths, 2010) focus on user comments in Tianya. HAN Ruixia (HAN Ruixia, 2010) stated the basic conception and concluded the basic feature of microblog platform. ANG Rui (WANG Rui & JIN Yong sheng, 2010) explained the relationships of user's friend numbers and the degree of user's population, but did not introduce the data source through his article. Other researchers conducted the research in the measurement of user's influence, the exploration of users' relationship and the information propagation way (Ma Jun et al, 2013; Yuan Fuyong et al, 2010; Lian Jie, 2011). Very few studies analyze the data acquisition methods and the network characteristics of social network users group. In this study, we will compare the existing Network data acquisition method, and put forward an efficient and feasible social network data acquisition technology. At the same time, we will use the related theory of Webometrics and complex social network to analyze the group feature of

Sina Microblog. The research may provide a reference to the organizations who want to use the SNS for business promotion and other activities.

Data Acquisition

This paper intends to use the API (Application Programming Interface) and the web spider based on page parsing data acquiring strategy. Control the calling method and frequency of API through the program. Acquire and parse JSON object to realize an effective data acquiring.

The data acquisition scheme based on page extracting can achieve maximum data. However, it is very complicate to extract the effective information (M.Spiliopoulou, 2000; Zhou Lizhu & Lin Ling, 2005). The data fetching strategy based on API has a high performance, it is very convenient to acquire and parse the data on Sina Microblog through calling the API, such as User ID， birthdate， register time， friends/fans numbers. But Microsoft service provider will not open their whole API to users. Therefore, using the open API may only solve a little part of problems in microblog data acquisition. For example, in Sina Microblog, some API which has important inquiring function is not opened, as for the usable open API, there exist inquiring quantity limitation. In consequence, to get more usable data, we combine the web spider and page parsing other than API. The API interfaces in Sina Microblog below are provided in table 1.

Table 1. API Interface in Sina Microblog

Microblog Interface	Comment Interface	User Interface	Top Microblog Interface
Relationship Interface	Buddy Group Interface	ID Interface	Collection Interface
Topic Interface	Microblog Label Interface	User Label Interface	Register Interface
Searching Interface	Recommend Interface	Short Chain Interface	Message Interface
Public Service Interface	Location Service Interface	Social TV Interface	Geography Information Interface
Video Upload Interface	OAuth 2.0 Authority Interface	Map Engine Interface	Pay Interface

Because of Sina Microblog’s limitation of User Access in API, we call the API every 5 seconds and make the program automatic into sleep state when achieving the upper limit. To prevent the repeat grab of data induced by the interruption in fetching processing, we store the user ID to be fetching into the queue, every time we grab the user information, the program read the ID, and every time the grab is over, the ID will be deleted. Through this way, in data fetching processing, even though we need to interrupt and restart again and again, the data will not repeat.

Through the above methods, we obtain 635029 user relationship data and 7688 user data.

Data Analysis

Microblogging network belongs to the complex network, so some theories of complex networks can be used in the study of the topological properties of microblogging network. We will explore its degree distribution, the network topology relation and the characteristic path length through MATLAB. These three programs are helpful to our study as follows:

i. In-degree distribution function: It calculates the number of in-degree distribution (K) per user, then figures out possibility of every number P(K); Out-degree distribution function: The number of out-degree distribution per user will be counted, so does the possibility of each number.

ii. The network topology relation is a useful tool to describe relations among nodes.

iii. Function of characteristic path length: It can measure the average shortest path length in network on the basis of Floyd algorithm (Sang Hoon Lee et al, 2006).

The network degree distribution

In Sina Microblog, the relationship between users is bidirectional, rather than unidirectional. One can become a fan of others by concern someone at any time, and may become an idol of somebody in return. We introduce in-degree and out-degree to measure these relationships. In-degree is used to count the number of people who focus on the user, and out-degree is the number of people that the user is concerned about (Wang Lin & Dai Guanzhong, 2006).

Figure 1 shows the change of probability distribution of in-degree(K). There appears to be a strong downward trend in the series over the entire sample. In fact, between 0 and 20, the possibility of K grew locally. The maximum of P(K) appears at K=20, then the posibility seems to decrease fairly rapidly since 20, which means that most of users in Sina Microblog are concerned by small number of people. When K tends to 1000, the possibility is nearly close to zero, reflecting a fact that there are just a few of stars in Sina Microblog.

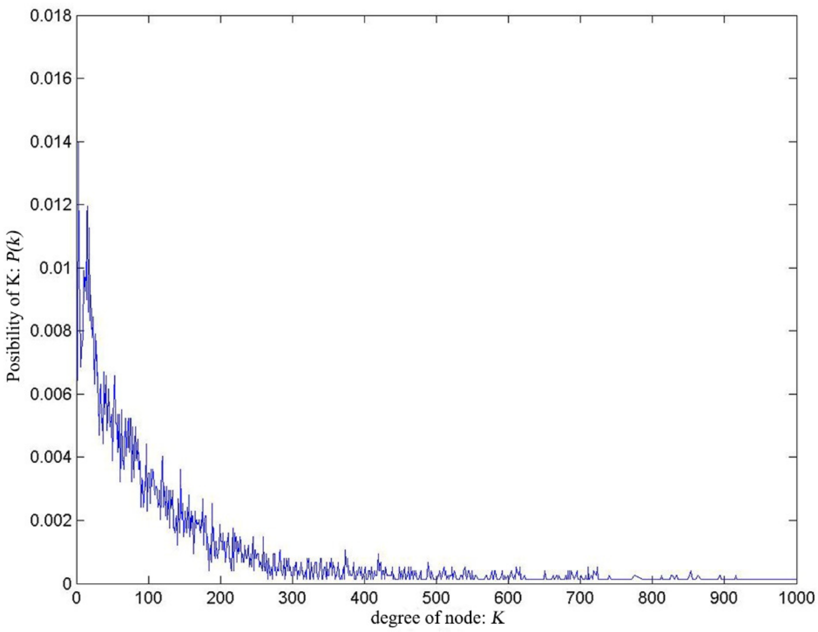


Figure 1. In-degree probability distribution

The curve was fitted in order to describe the feature of the in-degree probability distribution curve. The fitting curve of in-degree probability distribution is shown in figure 2. The probability distribution function is $P(\mathrm{K}) = \mathrm{a}_1\mathrm{K}^{-b1}$, and the fitting parameters are a1=0.09, b1=0.67. Then, natural logarithm is used on both sides $\ln P(\mathrm{K}) = \ln \mathrm{a}_1 - \mathrm{b}_1 \ln(\mathrm{K})$. We can get a negative linear which has a slope of b1 in double logarithmic coordinates (seen in Fig.3). The result is in accordance with characteristics of power-law distribution.

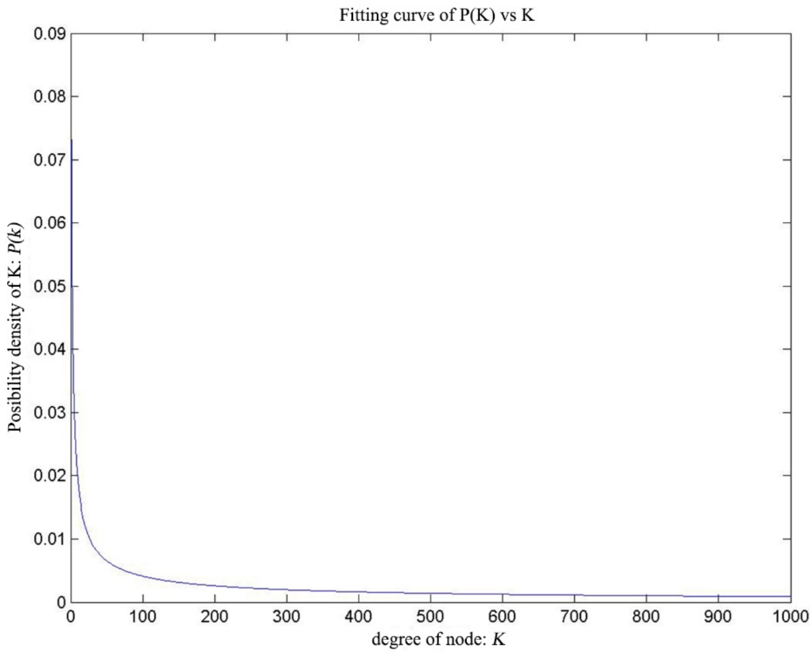


Figure 2. Fitting curve of in-degree probability distribution

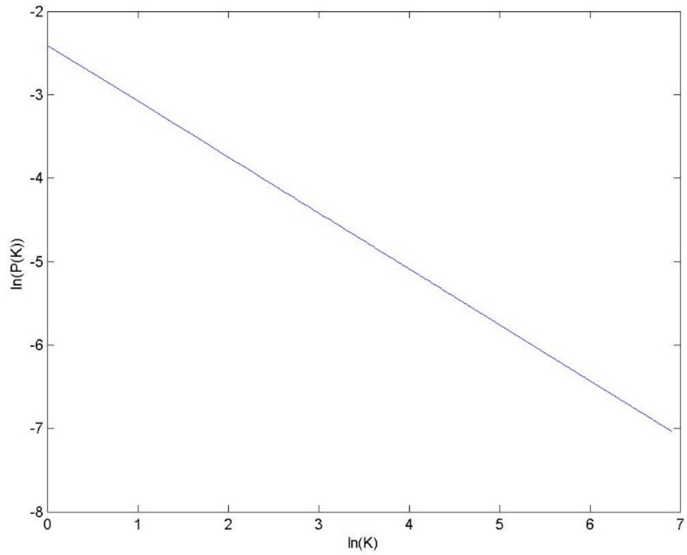


Figure 3. Correlativity between lnk and lnP (k) of in-degree

Figure 4 shows the probability distribution of out-degree(K). There appears to be a strong downward trend in the series over the entire sample. In fact, between 0 and 50, the possibility of K grew locally. The maximum of P(K) appears at K=50, then the possibility seems to decrease fairly rapidly since 50. When K tends to 2300, the possibility is nearly close to zero.

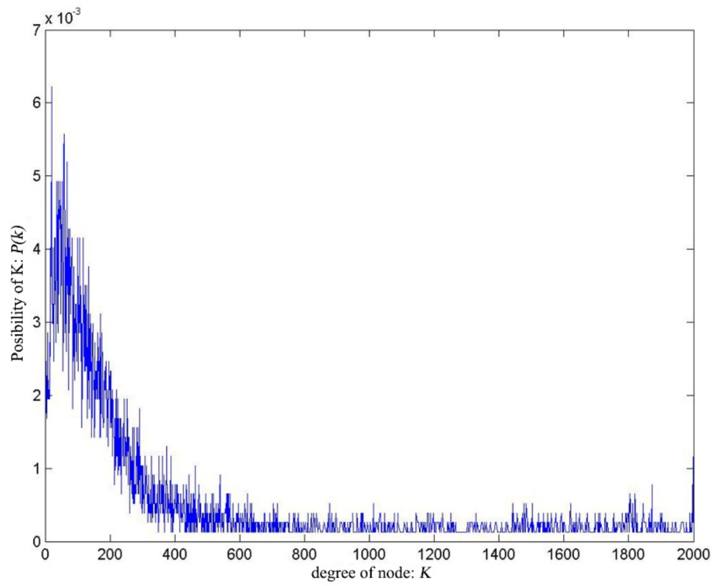


Figure 4. Out-degree probability distribution

The curve was fitted in order to describe the feature of the out-degree probability distribution curve, the out-degree probability distribution curve is shown in figure 5. The probability distribution function is $P(\mathrm{K}) = \mathrm{a}_2 \mathrm{K}^{-b_2}$, in which the fitting parameters are a_2=0.09, b_2=0.67. Then, natural logarithm is used on both sides $\ln P(\mathrm{K}) = \ln \mathrm{a}_2 - \mathrm{b}_2 \ln(\mathrm{K})$. We can get a negative linear, which has a slope of b2, in double logarithmic coordinates (shown in Fig.6). The result is in accordance with characteristics of power-law distribution, too.

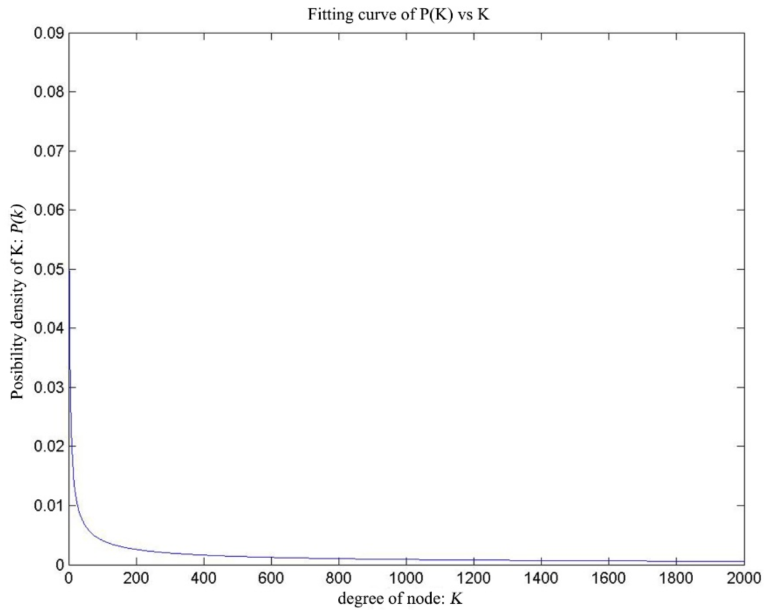


Figure 5. Fitting curve of Out-degree distribution

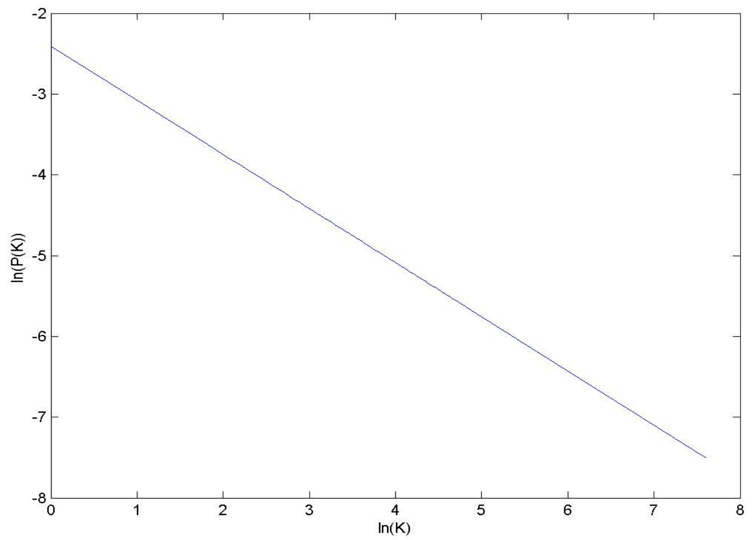


Figure 6. Correlativity between lnk and lnP (k) of out-degree

Network topology

Network topology reflects the connections among nodes. We use the topology function in MATLAB to process the data, and the result is shown in Fig. 7. There are a few nodes that connect with a lot of users. Most of nodes, meanwhile, just have a small cluster of connection line.

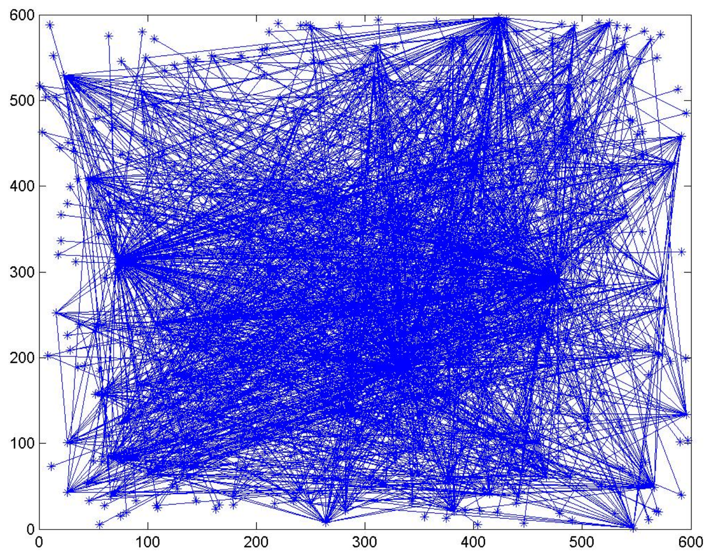


Figure 7. Network topology of Sina Microblogging

In order to make the connection more clearly, 300 nodes are randomly selected and analyzed as seen in Fig. 8. In the figure, the nodes marked 1,2,3 have the most connections. They are the central nodes or Star Junctions, which play an important role of the network. There are so many nodes connecting with them, so the information will spread very fast through the Star Junctions.

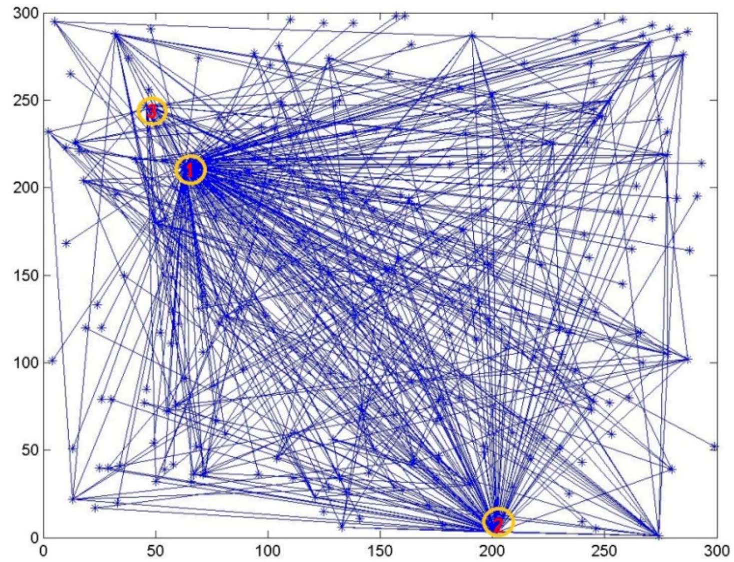


Figure 8. The local topology of 300 nodes

Characteristic Path Length

The characteristic path length is one of the most important and frequently-invoked characteristics of a social network. The path length of the two nodes is defined as minimum number of edges which can connect the two nodes. And the characteristic path length is the average of all the path lengths in the network.

The Characteristic Path Length of our obtained data is 3.7315, calculated by MATLAB. It means that the Sina Microblog is a Small World.

Conclusion

We introduce a method combining API with the web spider to fetch data from Sina Microblog, making data acquisition more easily and efficiently. Some theories of complex network are used to analysis the group characteristics of Sina Microblog. It shows that both in-degree and out-degree are small for most users. And we find that the characteristic path length is 3.7315, which means that the network of Sina Microblog is a Small World with a few of Star Junctions in it. We believe that the information is of value for applications based on Sina Microblog.

References

Westman S, Freund L. (2010). Information interaction in 140 characters or less: Genres on twitter//IIiX2010—Proceedings of the 2010 Information Interact ion in Context Symposium. New Brunswick, USA: *Association for Computing Machinery*, 2010:323 - 326.

Ye Wu,Jurgen Kurths. (2010). Human comment dynamics in on-line social systems.458th WE-Heraeus-Seminar SYNCLINE 2010: Synchronization in Complex Networks.

HAN Ruixia. (2010). The influence of microblogging on personal public participation //Proceedings of the 2010 IEEE 2nd Symposium on Web Society, SWS 2010. Beijing, China: *Association for Computing Machinery*, 2010:615-618.

WANG Rui, JIN Yong sheng. (2010). An empirical study on the relationship between the followers number and influence of microblogging //Proceedings of the International Conference on E-Business and E-Government, ICEE 2010. Guangzhou, China: *IEEE Computer Society*, 2010: 2014-2017.

Ma Jun, Zhou Gang, Xu Bin, Huang Yongzhong. (2013). Analysis of user influence in microblog based on individual attribute features. *Application Research of Computers*, 30(8):2483-2487.

Yuan Fuyong, Feng Jing, Fu Qianqian. (2012). Influence Index Model of Micro-blog User. *New Technology of Library and Information Service*, 6:60-64.

Lian Jie, Zhou Xin, Cao Wei, Liu Yun. (2011). SINA microblog data retrieval. *J T sing hua Un iv (Sci & Tech)*, 51(10):1300-1305.

M.Spiliopoulou. (2000). Web Usage Mining for Web Site Evaluation. *Communications of ACM*, 43(8):127-134.

Zhou Lizhu, Lin Ling. (2005). Survey on the research of focused crawling techn ique. *Computer App lications*, 25(9):1965-1969.

Sang Hoon Lee, Pan-Jun Kim, Hawoong Jeong. (2006). Statistical proerties of sampled networks. *The American Physical Society*, Phys.Rev.E,73:016102.

Wang Lin, Dai Guanzhong. (2006). The degree distribution of complex networks. *Journal of Northwestern Poly technical University*, 24(4):405-408.

Survey of the Editorial Board Members for Journals of Library and Information Science in China

Jiang Chunlin, Zhang Liwei and Liu Xue

WISE Lab, School of Public Administration and Law, Dalian University of Technology, Dalian 116085 China
chunlinj7873@163.com, taiyiliwei@126.com, 314014155@qq.com

Abstract

The editors of journal are the gatekeeper to hold the quality and direction of journals. Normally the editors of the journal are served as the "academic authority" in the corresponding discipline. Due to the limitation of the academic authority, it's common that one expert takes the editor of many journals, which is the interlocking editorship phenomenon. Base on this, we firstly built a social network matrix of the editorial board members about CSSCI journals of Library and Information Science. After that, we utilized the K-cores analysis of SNA to separate the network structure and to find the core subgroup, used the centrality analysis to verify the core of the core subgroup nodes. Secondly, we count the number of published papers in the past five years, the total cites and total downloads about this papers, and H index of total editorial board members. All of this data of bibliometrics reflects the academic performance of the editorial board member in the past five years, and then we proof the academic quality of journals through the data. At last, we discuss some problems which exist in the current regime.

Keywords: editorial board members; SNA; K-cores analysis; academic quality of journals; impact factor

Introduction

The academic journals are the major carrier to report the research achievement and spread the knowledge. Also they promote the development of academic discipline, personnel training, searching management and information dissemination. The editorial board is one of the important functional organizations, it's the gatekeeper to hold the direction and quality of journals, and it's responsible for the direction and quality of journals. Normally the functions of the editorial board are defined as the following: firstly, suggesting and deciding the editorial policy and direction of development about journals; secondly, promoting the development of journals; thirdly, encouraging others to contribute the publication; fourthly, reviewing the manuscripts and recommending the appropriate reviewer (Hames, 2001). For this, how to fully play the role of editorial board is very important to improve the quality of journals.

Usually the members of editorial board are the prominent scholars and the leaders of academic authority with high academic level, and if editorial board could work well depending on the good working members of editorial board (Baccini & Barabesi, 2011). Due to the limitation of the prominent scholars and academic authority, it's a common phenomenon that one expert takes the duty of multiple journals' editorial board, which is called 'interlocking editorship' (Baccini & Barabesi, 2010). As a result, there may be existed a core subgroup which contain some editorial board members taking the duty of multiple journals' editorial board among all the journals' editorial board members. Meanwhile, editorial board consisted by various famous experts in a particular academic fields, who have a direct influence on the exertion of the functions of the editorial board. So the experts' academic competence has a direct influence on journals' academic quality. Based on this, this paper taking journals of Library and Information Science included in CSSCI as example, uses the K-cores analysis and centrality analysis of SNA (Caroline, 2000; Newman, 2001) to find the core subgroup of all the journals' editorial board members. Meanwhile, this paper collects various kind bibliometrics datas of all the journals' editorial board members to verify the journals' academic quality, taking this to examine if the editorial board members take full duty.

Data collection and processing

The data of this paper collected from the editorial board member's information of the 23 journals (included the extended version) of Library and Information Science which are indexed by the Chinese Social Sciences Citation Index (CSSCI) which was published in 2012 ~ 2013. CSSCI is developed by the social sciences research evaluation center of the Nanjing University; it's used to retrieve data of the publications in the Chinese social science. The quality of the journals indexed by CSSCI possesses certain objectivity and authority. In the CSSCI, there totally are 23 journals included in the field of Library and Information Science (2012~2013): 18 core journals (the Archives Science Study and Archives Science Bulletin in Archival science field are not included) and 5 journals in the extended part of the CSSCI, the details are shown in table 1.

Tab. 1 Journals of Library and Information Science included in CSSCI (2012~2013)

number	name of the journals	number	name of the journals
1	Journal of Library Science In China	13	Library Tribune
2	Journal of Academic Libraries	14	Journal of Information
3	Journal of The China Society For Scientific and Technical Information	15	Library
4	Library and Information Service	16	New Technology of Library and Information Service
5	Document,Information & Knowledge	17	Library Work and Study
6	Information Studies:Theory & Application	18	Library Theory and Practice
7	Library and Information	1'	Researches In Library Science（extended）
8	Journal of The National Library of China	2'	Modern Information（extended）
9	Information Science	3'	New Century Library（extended）
10	Library Development	4'	Library Work in Colleges and Universities（extended）
11	Information and Documentation Services	5'	Journal of the Library Science Society of Sichuan（extended）
12	Library Journal		

After collecting the data, it is totally received 523 experts who hold the job of 23 kinds of journal editorial board. Due to some editorial board members taking the job of multiple journals' editorial board (interlocking editorship phenomenon), there are some editorial board members are the same person. Subtracted the duplicate, it is received 308 experts at last.

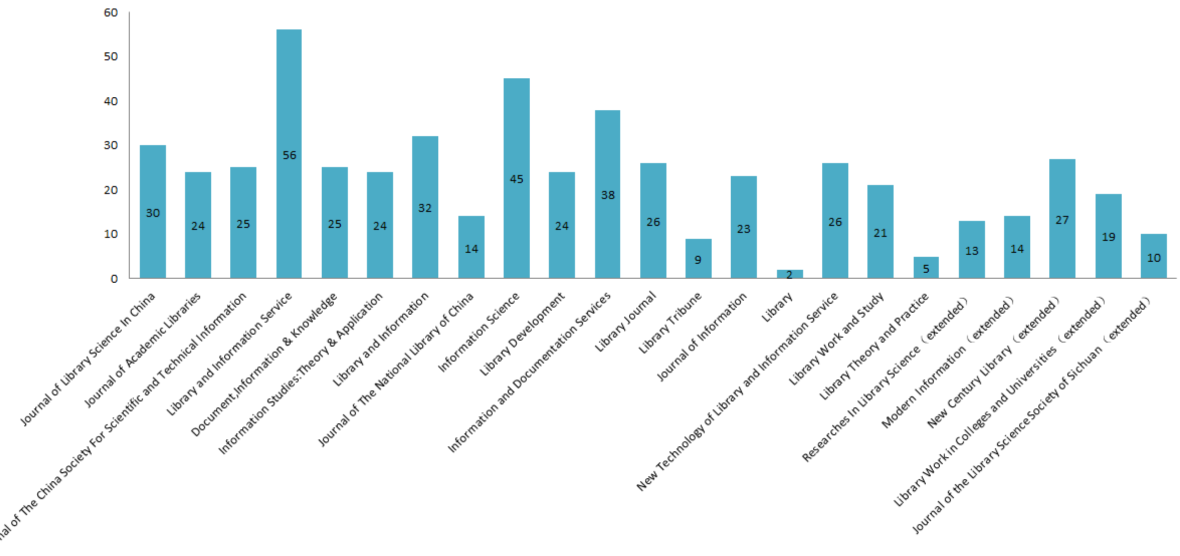


Fig 1. The Number of Editorial Board of Journals

Social Network Analysis about editorial board members

Network analysis

Due to the existence of 'interlocking editorship' phenomenon, we process the collected data of editorial board members and built a co-occurrence matrix, the details shown in table 2. The co-occurrence of editorial board members stands for the relationship between the members. During the data process, we develop custom java code to process the collected data, with the custom code, the data is converted to the format of Pajek network which could be identified by Ucinet.

Tab. 2 The Social Network Matrix of Editorial Board Members (in part)

	Ma	Chen	Cheng	Fan	Wang	Ke	Leng	Li	Ma	Wu
Ma Feicheng		5	3	4	2	3	3		3	3
Chen Chuanfu	5		6	7	6	7	5	4	5	4
Cheng Huanwen	3	6		8	6	6	3	4	4	2
Fan Bingsi	4	7	8		6	6	4	5	5	2
Wang Yuguang	2	6	6	6		5	3	3	3	2
Ke Ping	3	7	6	6	5		3	4	5	2
Leng Fuhai	3	5	3	4	3	3		3	4	4
Li Guoxin		4	4	5	3	4	3		3	
Ma Haiqun	3	5	4	5	3	5	4	3		3
Wu Yishan	3	4	2	2	2	2	4		3	

Ucinet is one of widely applied social network analysis software, it's famous with the advantage of easily to use and high data compatible ability, and it could be analyzed by Netdraw to get the visualization analysis results. After inputting the collected data to ucinet, we also could use Netdraw to get the co-occurrence network mapping. The on-line relations set as ≥2, which is mean the times of co-occurrence ≥2, we get the co-occurrence network mapping composed by 73 nodes, as shown in figure 2.

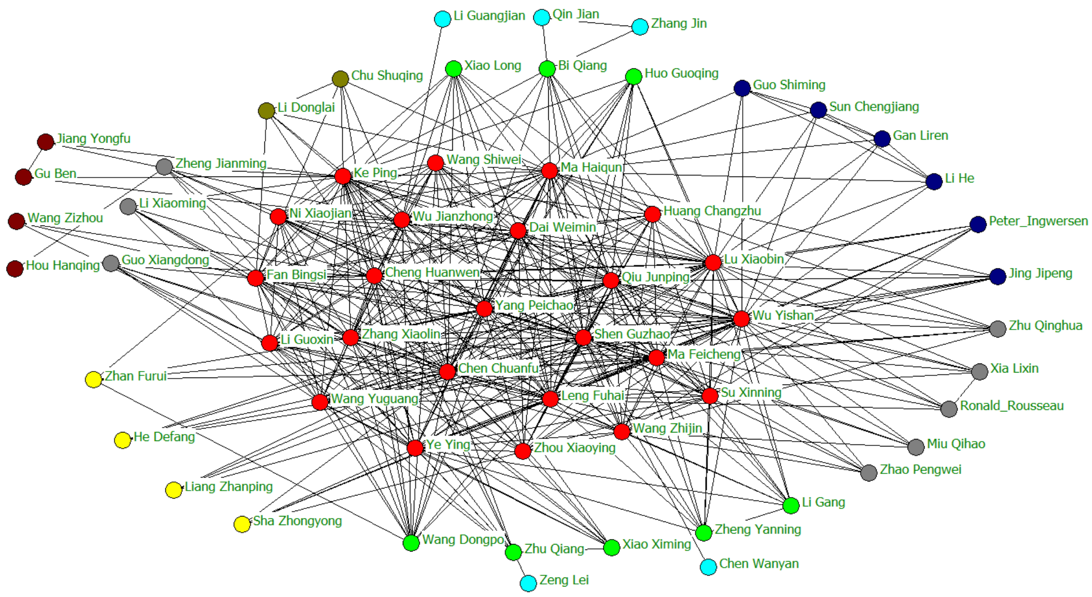


Fig 2 .The Co-occurrence Network Mapping of Editorial Board Members

In the graph above, every note stands for an editorial board member, and the lines represent that two notes exist co-occurrence relations. The color of notes stands for the value of K-cores (Baxter, 2012; Newman, 2003), which represent the degree of core of the whole network

(Zhang & Zhao et al, 2007). In Figure 2, the red notes are the 14 value of K-cores, and they are largest value of notes. So the red notes are the core notes of the whole network.

Centrality analysis

Centrality analysis is one of the most important methods of SNA. The note's degree of centrality stands for the degree of core in the network. The centrality reveals the power and status of the individual or organization in the social network which is stand by the node. Centrality analysis could divide into 'degree analysis', 'betweenness analysis', 'closeness analysis'. As is shown in Tab 3, we can precisely get the centrality data of every note and the mean centrality data of all the notes in the network.

Tab 3. The Analysis of Node Centricity (in part)

number	member	Degree	NrmDegree %	Betweenness	nBetweenness %	Farness	nCloseness
1	Ma Feicheng	97	16.840	89.666	3.508	110	65.455
2	Chen Chuanfu	161	27.951	244.216	9.555	96	75.000
3	Chen Huanwen	112	19.444	86.583	3.387	111	64.865
4	Fan Bingsi	130	22.569	106.556	4.169	108	66.667
5	Wang Yuguang	89	15.451	42.248	1.653	116	62.069
6	Ke Ping	130	22.569	201.376	7.879	104	69.231
7	Leng Fuhai	106	18.403	90.147	3.527	111	64.865
8	Li Guoxin	67	11.632	32.649	1.277	123	58.537
9	Ma Haixian	112	19.444	181.397	7.097	107	67.290
10	Wu Yishan	92	15.972	140.825	5.510	112	64.286
11	Shen Guzhao	120	20.833	224.773	8.794	102	70.588
12	Zhang Xiaolin	109	18.924	74.593	2.918	110	65.455
13	Wu Jianzhong	70	12.153	114.374	4.475	120	60.000
14	Yang Peichao	76	13.194	26.017	1.018	118	61.017
15	Ye Ying	74	12.847	73.120	2.861	116	62.069
16	Lu Xiaobin	107	18.576	180.477	7.061	105	68.571
17	Qiu Junping	119	20.660	135.855	5.315	104	69.231
…	…	…	…	…	…	…	…
Mean	n/73	46.874	1.902	35.192	1.377	142.384	52.189

K-cores analysis is used to reveal the hierarchy attribute of the structure of the social network, to identify the distribution of the subgroup in the whole network, and to find the core subgroup of the network (Wettler & Rapp,1993; Seidman,1983; Bollabás,1984). Though K-cores analysis and centrality analysis, we could receive a topological network structure with out-to-in of the whole editorial board members of Library and Information Science included in CSSCI. In the network mapping, it is emerged some notes with high value of K-cores, as 'Qiu Junping', 'Leng Fuhai', 'Ke Ping', 'Ma Haiqun', 'Fan Bingsi', 'Wu Yishan', 'Shen Guzhao' and so on. They are the core notes of the whole network. As is shown in Tab 3, they are also the highest centrality data of all the notes in the network. Thus, it is demonstrated they are the core notes of the whole network. As a result, such red notes as 'Qiu Junping', 'Leng Fuhai', 'Ke Ping', 'Wu Yishan', 'Shen Guzhao' etc. shown in figure 2 constitute the core subgroup of the whole network.

Academic performance of the editorial board members

In order to have an acquaintance with the editorial board members' academic performance, we collected some data of bibliometrics which could reflect the academic performance of the whole editorial board members, including published papers in the last five years, cited numbers of published papers, H Index etc. of the total 308 editorial board members (shown in Tab 4.).

Tab 4. The Statistics of Bibliometric indicators for the Editor Board Member (in part)

Num	editor board member	published papers（piece）	editor board member	cited numbers (times)	editor board member	download numbers (times)	editor board member	H Index (-)
1	Qiu Junping	210	Qiu Junping	818	Qiu Junping	70267	Qiu Junping	39
2	Wang Zhijin	127	Fan Bingsi	536	Zhu Qinghua	47416	Jiang Yongfu	39
3	Wu Yishan	124	Zhu Qinghua	517	Zhang Xiaolin	33190	Zhang Xiaolin	38
4	Pan Yuntao	109	Liu Wei	514	Ke Ping	29791	Ma Feicheng	30
5	Zhu Qinghua	108	Zhang Xiaolin	465	Wang Zhijin	29159	Huo Guoqing	30
6	Ma Haiqun	103	Ke Ping	457	Xiao Ximing	24579	Ke Ping	29
7	Dai Tao	96	Wu Yishan	418	Chu Jingli	23146	Xiao Ximing	29
8	Zheng Yanning	80	Pan Yuntao	374	Wu Yishan	22898	Zhu Qinhua	28
9	Bi Qiang	77	Ma Haiqun	359	Ma Haiqun	22753	Fan Bingsi	28
10	Ke Ping	75	Xiao Ximing	330	Pan Yuntao	22602	Wang Zizhou	28
11	Leng Fuhai	71	Chu Jingli	319	Ma Feicheng	22270	Chen Chuanfu	27
12	Xiao Ximing	69	Wang Zhijin	294	Ye Jiyuan	19231	Wu Weici	27
13	Fu Rongxian	61	Jiang Bifu	287	Bi Qiang	19007	Wang Shiwei	26
14	Ma Feicheng	59	Bi Qiang	255	Fan Bingsi	18509	Li Guoxin	26
15	Zhang Zhiqiang	58	Chen Chuanfu	254	Huo Guoqing	15432	Ma Haiqun	25
16	Zheng Jianming	57	Wang Zizhou	253	Zheng Jianming	15350	Hu Changping	25
17	Huo Guoqing	57	Li Guoxin	252	Zheng Yanniing	15318	Chu Jingli	24
18	Su Xinnin	53	Sun Tan	252	Dai Tao	15162	Liu Ziheng	24
19	Zhang Xiaolin	50	Jing Jipeng	251	Lai Maosheng	14893	Wang Zhijin	23
20	Zeng Jianxun	49	Ye Jiyuan	247	Leng Fuhai	14860	Cheng Huanwen	23
21	Zhang Zhixiong	49	Wang Shiwei	244	Hu Changping	13710	Zheng Jianming	22
22	Ye Jiyuan	47	Dai Tao	242	Sun Jianjun	13595	Jing Jipeng	22
…	…	…	…	…	…	…	…	…
261	—	0	—	0	—	0	—	0
…	…	…	…	…	…	…	…	…
308	—	0	—	0	—	0	—	0

*1 the number of published papers is derived from CNKI(China National Knowledge Infrastructure); the number of published papers and the cited numbers are derived from CNKI and Wanfang database.

*2 the time period for the number of published papers, the cited numbers and the downloaded numbers of collected papers is 2009.02~2014.0, the collected time period for H index is unlimited.

*3 the statistics of the cited numbers contains the number of cited by self and by others.

These bibliometrics indicators could reveal the status of academic activity and the academic influence of the experts. The number of published papers, total cited number, and total download numbers directly reflects the academic ability and influence of the experts. But the distribution of quantitative value is not irrationality. For example, there are some experts with zero published papers or total cited number etc. shown in Tab 4, from which we can find the problems which exist behind of the phenomenon. After collecting the data of academic performances of 308 editorial board members, we had to sum the data of all the editorial board members according to the list of a particular journal.

Tab 5. Bibliometric Data of the Journal's EBM (editorial board members)

Name of Journal	Number of Journal EBMs	Impact Factor of the journal	Total number of the published papers by the Journal's EBMs	Total cited of the published papers by the Journal's EBMs	Total downloaded of the published papers by the Journal's EBMs	Total H index of the Journal's EBMs
Journal of Library Science In China	29	3.238	759	3823	246918	438
Journal of Academic Libraries	25	1.778	649	4229	250080	441

JCSSTI (*1)	24	1.003	1270	4032	320980	352
Library and Information Service	56	0.808	1703	7422	491658	770
Document,Information & Knowledge	24	0.975	976	4562	322526	438
ISTA (*2)	24	0.897	9	13	2178	17
Library and Information	26	1.183	781	3559	218952	377
JNLC (*3)	14	1.364	301	1689	109406	213
Information Science	44	0.783	1719	6147	482453	677
Library Development	24	0.959	599	3790	191197	365
IDC (*4)	38	0.886	1021	4140	302781	430
Library Journal	26	1.045	334	1953	97109	233
Library Tribune	2	0.878	3	2	281	5
Journal of Information	23	0.806	901	3330	265539	309
Library	2	0.665	7	4	632	4
NTLIS (*5)	26	0.789	565	2648	163019	329
Library Work and Study	20	0.697	337	1451	100341	188
Library Theory and Practice	5	0.521	26	26	1686	17
Researches In Library Science	13	0.789	65	99	7167	61
Modern Information	14	0.482	522	1953	140393	160
New Century Library	27	0.356	335	993	75678	229
LWCU (*6)	19	0.490	683	3430	241156	338
JLSSS(*7)	10	0.390	124	670	47009	135
Correlation Index	-	-	0.152	0.274	0.213	0.286

NOTE:

*1, JCSSTI: 'Journal of The China Society For Scientific and Technical Information';

*2, ISTA: 'Information Studies: Theory & Application'; *3, JNLC: 'Journal of The National Library of China';

*4, IDC: 'Information and Documentation Services'; *5, NTLIS: 'New Technology of Library and Information Service';

*6, LWCU: 'Library Work in Colleges and Universities'; *7, JLSSS: 'Journal of the Library Science Society of Sichuan'.

*8, Impact Factors of the journals come from CNKI in 2013.

As is shown in the above table, the total number of the published papers of the Journal refers to the sum total published papers in the last five years by all editorial board members of the journal. Likewise, the total cited of the Journal and the total downloaded of the Journal refers to the sum total cited number and downloaded number of the papers which was published by the all editorial board members of the journal in the last five years. And the total H index numerical value refers to the sum total value of H index of all editorial board members of the journal. All the data of the journal of above could be able of reflect the academic performance and the influence of academic achievements of all the journal's editorial board members. As the individual editorial board member's quantitative value, the total quantitative value of editorial board members of a particular journal also was not irrationally distributing. After that, we had carried on a correlation analysis between the academic performances of editorial board members and impact factor which stands for the academic quality of journals.

After calculation, we get the correlation coefficients between impact factor of the journals and the total number of the published papers by the Journal's EBMs, the total cited of the published papers by the Journal's EBMs, the total downloaded of the published papers by the Journal's EBMs, the total H index of the Journal's EBMs. What the results are '0.152, 0.274, 0.213, 0.286', as is shown in the above Tab 5. The numerical values of correlation coefficients reflects that there are positive correlations between the data of bibliometrics of the journals' editorial board members and impact factor of the journals, but the degree of correlation is very weak.

Results and conclusion

With the above analysis, we get the following conclusions.

Firstly, with the K-cores analysis to process the collecting data of the editorial board members which was extracted from CSSCI journals, we find the subgroup in the co-occurrence network of all the editorial board members. Meanwhile, we demonstrated the nodes included in subgroup have strong degree of core through the centrality analysis. To find the core group which leads the development of the Library and Information Science is the main purpose of K-cores analysis and centrality analysis.

Secondly, from the statistical data of bibliometrics of editorial board members, this paper gets fully cognition of the academic performance of all the editorial board members. The research finds out that most of editorial board members would be positive to take part in the activities of scientific research, they have been published many papers and have been produced specific academic impact. But there are some other editorial board members have not published any papers, also the correlation coefficients between impact factor of the journals and the total number of the published papers by the Journal's editorial board members are very weak. Which the two consequences reflect some problems existing in the current regime of editorial board in China, one of important part is that editorial board members of journal don't make full commitment to the job. Take the following reasons may account for the questions. Firstly, the duty of editorial board for journals is not the major job for many expert, most of them take the part-time job for the journals, because they have many social affairs and the time is limited. For another reason, some editorial board members are too old to take the job of editorial board. Thirdly, many experts may take the job for many journals, which will also influence the function of the editorial board.

References

Hames I. (2001) . Editorial boards: realizing their poential. *Learned Publishing* ,14 (4) :247-256.

Baccini A, Barabesi L. (2011). Seats at the table: the network of the editorial boards in information and library science[J]. *Journal of Informetrics*, 5(3): 382-391.

Baccini A, Barabesi L.(2010). Interlocking editorship. A network analysis of the links between economic journals[J]. *Scientometrics*, 82(2): 365-389.

Caroline S Wagner, LoetLeydesdorff. (2005). Mapping the network of global science: comparing international co-authorships from 1990 to 2000[J]. *International Journal of Technology and Globalisation*, 1(2), 185-208.

Newman, M. E J. (2001). The structure of scientific collaboration networks[J]. *Proceedings of the National Academy of Sciences*, 98(2), 404-409.

Baxter G J, Dorogovtsev S N, Goltsev A V, et al. (2012). *K-Core Organization in Complex Networks[M]//Handbook of Optimization in Complex Networks*. Springer US, 2012: 229-252.

Newman M E J. (2003). The structure and function of complex networks[J]. *SIAM review*, 45(2): 167-256.

Zhang H H, Zhao H, Cai W, et al. (2007) . A qualitative method for analysis the structure of software systems based on k-core[J]. *Special Issue on Software Engineering and Complex Networks of Dynamics of Continuous, Discrete and Impulsive Systems Series B*, 14(S6): 18-24.

Wettler M, Rapp R. (1993). Computation of word associations based on the co-occurrences of words in large corpora[C]//*Proceedings of the 1st Workshop on Very Large Corpora*. 1993: 84-93.

Seidman S B. (1983). Network structure and minimum degree[J]. *Social networks*, 5(3): 269-287.

Bollabás B. (1984). The evolution of sparse graphs[C]//*Graph Theory and Combinatorics: Proc. Cambridge Combinatorial Conf. in honor of Paul Erdös* (B. Bollabás, Ed.). 1984: 35-57.

The Effect of Geographical Proximity on Organizational Knowledge Sharing

Leila Nemati-Anaraki* and Roya Pournaghi**

*Iran University of Medical Sciences, Tehran, Iran
lnemati@yahoo.com

**Assistant Professor of Iranian Research Institute for Information Science & Technology, Tehran, Iran
pournaghi@irandoc.ac.ir

Abstract

Today, knowledge is considered as the main driver in "knowledge-based economy". Increasing knowledge sharing would have a positive effect on organizational performance. There are a set of factors which effect on knowledge sharing such as: ICT supports, age, size, and geographical location or proximity. In this article we focus on the geographical proximity. The purpose of this paper is whether there is a significant relation among faculty members in Tehran and other cities in terms of the extent of inter- and intra-organization knowledge sharing and also how is knowledge sharing between faculty members in each cities of Iran?. The present study was based on mixed method and applied research. For data analysis, statistical software SPSS15 and two-sample mean T-student was used. Results indicates a significance difference among faculty members in Tehran and other cities in terms of the extent of intra-organization knowledge sharing and no significance difference was observed among them in terms of the extent of inter-organization knowledge sharing. With the comparison of the amount of intra- and inter-organizational knowledge sharing, Isfahan, Tehran, Tabriz, Shiraz, Hormozgan, and Baghyat-allah medical university have higher than the average knowledge sharing and in medical research centres Pastor, Ahvaz, Behzisti, have higher than the average knowledge sharing. By focusing on the results of this research policy makers can promote this process in all country universities.

Keywords: Proximity, Geographical proximity, Knowledge sharing, Medical Universities, Faculty Members, Iran.

Introduction

The proximity concept has captured a prominent position in the scientific literature dealing with intra- and inter-organizational collaboration, innovation, and regional economic development (Mackinnon *et al.,* 2002; Oerlemans *et al.,* 2001). It is an important emerging concept in several fields. When the proximity concept is used, what is often actually meant is geographical proximity. However, other forms of proximity, such as institutional proximity, organizational proximity, cultural proximity, social proximity, and technological proximity are used as well (Gill & Butler, 2003; Greunz, 2003; Meisters & Werker, 2004). Even though all of these dimensions of the concept of proximity refer to "being close to something measured on a certain dimension", they are certainly not identical. Various methods have been proposed for sharing knowledge in organizations in general and for transferring and sharing implicit knowledge in particular. The importance of geographical proximity for interaction and knowledge sharing has been discussed extensively in recent years. There is increasing consensus that geographical proximity is just one out of many types of proximities that might be relevant (Broekel & Boschma, 2012). Geographical proximity is the most significant and applicable method of sharing knowledge, which have received considerable attention in the related literature. One technique for designing an appropriate workplace in order to improve sharing knowledge in organizations is designing it in a way that it leads to creating shared cooperative settings near people's commuting or increasing the interaction of those who should be involved in sharing (Nieminen, 2005; Hau & Evangelista, 2007; van Wijk *et al.*, 2008; Seyyedeh *et al.*, 2009).

Geographical proximity integrates the social dimension of economic mechanisms, or what is sometimes called functional distance. In other words, the reference to natural and physical constraints is an important aspect of geographical proximity but other aspects are equally important in its definition: the aspect of social structures such as transport infrastructures that facilitate accessibility, or the financial mechanisms that allow the use of certain communication technologies. It is necessary to take this definition of proximity further by distinguishing permanent geographical proximity, which corresponds to the co-localization of firms, from temporary geographical proximity, which lies on momentary face to face interactions enabling actors to meet without necessarily requiring co-localization (Rallet & Torre, 2005; Mitchel *et al.*, 2010). The definition of this dimension of proximity differs slightly between different authors. Some studies look at the distance between two interaction organizations, whereas others look at the presence of groups of firms in a geographical unit. Nevertheless, the definitions of geographical proximity are all fairly similar and use the same underlying mechanism for explaining the importance of geographical proximity. The importance of geographical proximity in intra- and inter-organizational collaboration lies in the fact that small geographical distances facilitate face-to-face interactions (both planned and serendipitous) and, therefore, fosters knowledge transfer and innovation. The main reasoning behind these effects is that short geographical distances bring organizations together, favour interaction with a high level of information richness and facilitate the exchange of, especially tacit, knowledge between actors (Torre & Gilly, 2000). The larger the distance between actors, the more difficult it is to transfer these tacit forms of knowledge. This is even argued to be true for the exchange and use of codified knowledge, because its interpretation still requires tacit knowledge and thus spatial proximity (Howells, 2002; Lilleoere & Hansen, 2011).

The need for geographical proximity is generally not permanent. It affects certain phases of the interaction: the phase of negotiation in a transaction, the definition of the organizational framework and guidelines of cooperation, the realization of its initial phase in the case of a technological alliance, the necessity to share equipment in the experimental phase of a common research project or to exchange knowledge and above all to know personally the researchers belonging to a scientific community etc. Short or medium-term visits are then sufficient for the partners to exchange – during face to face meetings – the information needed for cooperation. As a result permanent co-localization is not necessary even for activities, where physical interaction plays an important role in the coordination (services co-produced by the provider and the user, knowledge-intensive activities such as innovation and R&D activities). This is what we call the need for temporary geographical proximity (Gallasud & Torre, 2004).

Based on the above, the goals of this paper is to identify significant difference among the extent of inter- and intra-organization knowledge sharing by faculty members of the universities and research centres affiliated with Ministry of Health Treatment and Medical Education in terms of geographical area where they work?, Compare the factors which might influence knowledge sharing among faculty members in universities and research centres, and Rank the universities and also research centres participating in the current study with regard to the extent of knowledge sharing and effective factors.

Literature Review

To gain insight into the different dimensions of proximity and their definitions, a literature review has been conducted. Geographical proximity, which is denoted as territorial, spatial, local or physical proximity as well, is the most frequently used dimension of proximity in the literature. Many studies do not even explicitly state that geographical proximity is being used, but just use the term “proximity”. Several authors have put forward the notion of temporary geographical proximity (Gallaud & Torre, 2005; Torre & Rallet, 2005). This notion implies that actors need not be in constant geographical proximity when collaborating, but that meetings, short visits and temporary co-location might be sufficient for actors to build other forms of proximity (such as organizational), which subsequently allow collaboration over large geographical distances. Development in communication technologies have made it feasible for actors to work together despite physical dispersion of group members. The study by Cramton (2001) focus on an experiment in which team members had to collaborate without meeting face-to-face and therefore had trouble building mutual knowledge. Cramton proceeds by linking certain problems in building mutual knowledge to the lack of face-to-face contacts and thereby attributes all problems in the organizational collaboration to the lack of geographical proximity. The problems she described are: failure to communicate and retain contextual information, unevenly distributed information, differences in the salience of information, relative differences in the speed of access to information and differences in interpreting the meaning of silence.

Although information technology is basically viewed as a tool for sharing knowledge, a large number of people have claimed that technologies can play the role of a driving force in sharing knowledge as well. The results of various studies indicate that the interaction among those who receive knowledge, physical proximity, and finally the type of shared knowledge would make an impact on how knowledge is shared (Jansen van Vuuren, 2011). Physical proximity is closely related to the extent of interaction and sharing knowledge. For instance, an individual operating in a floor interacts more than most other individuals in different floors of the same building. In the same way, the groups in various organizations or cities have lower extent of interaction.

In a number of studies, Knoben & Oerlemans (2006), Torre & Rallet (2005), and Gallaud & Rallet (2004) also concluded that physical and geographical proximity would result in more face-to-face communication and sharing more implicit knowledge and thereby, higher knowledge creation and creativity. Filippi & Torre (2003) also pointed to the significant relationship between geographical proximity and conducting joint projects in organizations located in an area.

Conversely, Gouza (2006) found out that geographical distance would not influence knowledge transfer among different organizations taking into account the existing communicative technologies. Nevertheless, it might be said that despite the available technologies, people are sometimes reluctant to share their knowledge and experience which might stem from lack of motivation. Morever, Xianyue & Rui (2013) in their research concluded that geographical and social proximity accelerates the transfer of knowledge, reduces the time and cost of knowledge search and overcomes the transferring barriers stemmed from the viscosity of knowledge. Janet *et al.* (2013) also indicated that knowledge sharing occurs through cross functionality, overlapping roles, and facilitated by close physical proximity in open workspaces; and knowledge reuse is often made tacitly, where common knowledge is prevalently embedded within the knowledge management processes of SMEs.

Research Methodology

The present study was based on mixed method and applied research which was conducted in 2012. The faculty members of medical universities and medical research centers supervised by the Ministry of Health Treatment and Medical Education of Iran constituted the population of the research. The statistic population consisted of 3430 persons employed as full-time faculty members at medical research centers and 12,428 persons employed as full-time faculty members of medical universities. Because the extent of the population, sampling techniques were used to collect the data needed to select the best sampling and according to the type of the universities and research centers, which "stratified random sampling" was used and a minimum required capacity of 423 people were provided. In order to perform sample capacity sufficiency, Bartlett and KMO's test were used (Table 1). The tool used in this study was questionnaire that face and content validity were confirmed by experienced professors. The Cronbach's alpha coefficient for this dimension was greater than 0.8, so, the reliability of the instrument was confirmed. For data analysis, statistical software SPSS15, and two-sample mean T-student was used.

Table 1. Bartlett and KMO's test

<table>
<tr><td>0/919</td><td colspan="2">Bartlett and KMO's test</td></tr>
<tr><td>32194/358</td><td>Chi-Square</td><td rowspan="3">Test</td></tr>
<tr><td>6786</td><td>df</td></tr>
<tr><td>0/000</td><td>Significant level</td></tr>
</table>

Results

Considering the first research goal "There is a significant difference among the extent of inter- and intra-organization knowledge sharing by faculty members of the universities and research centres affiliated with Ministry of Health Treatment and Medical Education in terms of geographical area where they work", the geographical area was divided into two categories of Tehran (the capital) and other cities. Then, two-sample mean t-student was conducted to compare the mean values of the extent of inter- and intra-organization knowledge sharing.

Table 2. The results of two-sample mean t-student for the significant difference between geographical area and inter- and intra-organization knowledge sharing

<table>
<tr><th colspan="2" rowspan="2"></th><th colspan="5">Test for the mean values</th><th colspan="2">Variance equality test</th></tr>
<tr><th>Significance level</th><th>df</th><th>T-test value</th><th>SD (Standard Deviation)</th><th>Mean</th><th>Significance level</th><th>value F</th></tr>
<tr><td rowspan="2">Intra-organization knowledge-sharing</td><td>Faculty members in Tehran</td><td rowspan="2">0/054</td><td rowspan="2">423</td><td rowspan="2">1/936</td><td>0/62945</td><td>3/2172</td><td rowspan="2">0/176</td><td rowspan="2">1/841</td></tr>
<tr><td>Faculty members in other cities</td><td>0/68951</td><td>3/0686</td></tr>
<tr><td rowspan="2">Inter-organization knowledge-sharing</td><td>Faculty members in Tehran</td><td rowspan="2">0/001</td><td rowspan="2">423</td><td rowspan="2">3/489</td><td>0/68336</td><td>2/3157</td><td rowspan="2">0/786</td><td rowspan="2">0/024</td></tr>
<tr><td>Faculty members in other cities</td><td>0/62925</td><td>2.0517</td></tr>
</table>

The results of two-sample mean t-student (Table 2) indicated a significance difference among faculty members in Tehran and other cities in terms of the extent of intra-organization knowledge sharing at the level of significance of 0.05. (Since the significance level of the test for the intra-organization knowledge sharing was higher than the test level, i.e. α=0/05 (level of significance=0.054 and df= 423). However, the average extent of sharing knowledge among faculty members in Tehran (M=3.21) was higher than that of their counterparts in other cities (M=3.06).

Moreover, according to Table 2, no significance difference was observed among faculty members in Tehran and other cities in terms of the extent of inter-organization knowledge sharing at the level of significance of 0.05. (Since the significance level of the test for the inter-organization knowledge sharing was lower than the test level, i.e. α=0/05 (level of significance= 0.001 and df= 423). Since the average extent of sharing knowledge among faculty members in Tehran (M=2.31) was higher than that of their counterparts in other cities (M=2.05), it might be concluded that faculty members in Tehran play a more significant role in inter-organization knowledge sharing than their counterparts in other cities.

Considering the second research goal, by comparing the factors which might influence knowledge sharing among faculty members in universities and research centres (participants of the current study). The findings revealed that considering the extent of sharing knowledge (inter- and intra- organization) among the faculty members of Medical Sciences Universities, Isfahan, Tehran, Tabriz, Shiraz, Baghiyat-Allah, Mashahd, Ahwaz, Kerman, Shahid Beheshti, Kasahn, Behzisti, and Yazd universities were at the level higher than average while Jahrom, Booshehr, Ghom, and Ilam universities were at the lowest level (Table 3).

Table 3. Ranking the universities participating in the current study with regard to the extent of knowledge sharing and effective factors

Name of University	Number	Sharing knowledge	Inter-organization knowledge-sharing	Intra-organization knowledge-sharing	Sub-structural factors (underlying factors)	Organizational factors	Individual (personal) factors
Isfahan	10	3/0050	2/5550	3/4550	2/8364	2/9711	3/6964
Tehran	37	2/9838	2/5324	3/4351	3/4963	3/2183	3/8948
Tabriz	21	2/9321	2/5333	3/3310	3/4329	3/2381	3/9473
Shiraz	15	2/9150	2/4933	3/3367	3/5273	3/3000	4/0048
Hormozgan	2	2/8375	2/2750	3/4000	2/0000	2/4079	2/8750
Baghyat-allah	4	2/7875	2/1375	3/4375	3/1136	3/0855	3/4554
Mashhad	36	2/7806	2/1750	3/3861	3/2298	3/0548	3/6329
Ahvaz	15	2/7800	2/2767	3/2833	3/3030	2/9684	3/9381
Kerman	21	2/7774	2/2238	3/3310	2/9437	2/7920	3/6667
Shahid Beheshti	13	2/7654	2/3462	3/1846	3/1678	2/9555	3/8462
Kashan	7	2/7179	2/0571	3/3786	2/4805	2/6992	3/4541
Behzisti	3	2/6583	1/9667	3/3500	2/2424	2/2982	3/5595
Yazd	13	2/6538	2/1383	3/1692	2/5594	2/7530	3/5495
Mean value of all universities	299	2/6196	2/1069	3/1323	2/8632	2/7946	3/5852
Kermanshah	4	2/5938	2/1500	3/0375	1/9318	2/4408	3/2411
Babol	4	2/5563	1/8750	3/2375	2/0455	2/1645	3/3036
Lorestan	8	2/5469	2/0250	3/0688	2/4091	2/6217	3/2634
Zanjan	2	2/4875	1/9000	3/0750	2/8182	3/0395	3/2500
Ardebil	5	2/4600	1/790	3/1300	2/2000	2/4474	3/1357
Uromieh	10	2/4325	1/8500	3/0150	2/2364	2/3842	3/6429
Tarbiat Modares	1	2/300	1/7000	2/9000	3/3636	2/5526	2/8929
Mazandaran	7	2/2500	1/7357	2/7643	2/3766	2/6466	2/9184
Zahedan	12	2/2042	1/7208	2/6875	2/4773	2/3136	3/1161
Ghazvin	7	2/2036	1/6000	2/8071	2/1169	2/1267	3/2092

Rafsanjan	6	2/1250	1/6167	2/6333	2/4848	2/7149	3/4405
Gonabad	1	2/0750	1/2000	2/9500	1/9091	2/1053	2/8214
Arak	2	2/0250	1/4000	2/6500	2/4545	2/1842	3/0893
Yasoj	6	1/9833	1/5250	2/4417	2/0303	2/0088	2/9524
Hamedan	7	1/9786	1/4857	2/4714	2/2078	2/2105	3/8061
Gilan	4	1/9500	1/4000	2/5000	2/3409	2/3618	2/9464
Shahre kord	2	1/8750	1/3750	2/3750	2/0909	2/3158	3/3214
Kordestan	3	1/8500	1/3333	2/3667	2/0000	2/1140	3/0952
Shahrood	4	1/8250	1/4375	2/2125	2/3864	2/5329	3/4286
Ilam	1	1/8000	1/2000	2/4000	1/9091	2/0789	3/3929
Ghom	3	1/7083	1/3833	2/0333	1/7879	1/7895	2/9286
Booshehr	2	1/5750	1/1500	2/0000	2/4091	1/9474	2/8571
Jahrom	1	1/5000	1/0000	2/0000	1/4545	1/5263	3/1071

The results showed that regarding the extent of sharing knowledge (inter- and intra-organization) among the faculty members of Medical Research centres, Pastoor, Ahwaz, Behzisti, Yazd, Mazandaran, and Tehran centres were at the level higher than average while those of Mashahd and Shahid Beheshti were at the lowest level (Table 4).

Table 4. Ranking the research centres participating in the current study with regard to the extent of knowledge sharing and effective factors.

Name of research centres	Number	Sharing knowledge	Inter-organization knowledge-sharing	Intra-organization knowledge-sharing	Sub-structural factors (underlying factors)	Organi-zational factors	Individual (personal) factors
Pastor	5	3/2100	2/8700	3/5500	2/8727	2/9947	3/3571
Ahvaz	6	3/0375	2/6167	3/4583	2/1970	2/2675	3/1429
Behzisti	4	2/9563	2/4875	3/4250	2/6136	2/6842	3/3036
Yazd	3	2/8083	2/2167	3/4000	2/7273	2/8158	3/7024
Mazandaran	7	2/7750	2/3000	3/2500	2/8571	2/7444	3/1224
Tehran	16	2/6391	2/1406	3/1375	2/3409	2/4622	3/0603
Mean value for the entire country	126	2/9540	2/1504	3/0377	2/5216	2/5102	3/2346
Baghyat-allah	14	2/5768	2/0964	3/0571	2/7727	2/6880	3/6505
Isfahan	5	2/5700	2/2300	2/9100	2/8364	2/6368	3/2571
Shiraz	21	2/5655	2/0905	3/0405	2/4199	2/4574	3/1293
Kerman	15	2/5350	2/0600	3/0100	2/5455	2/5667	3/1405
Yasoj	2	2/4875	2/0750	2/9000	2/5000	2/0000	3/3929
Tabriz	4	2/3625	1/8375	2/8875	2/1364	2/2763	3/0893
Gilan	3	2/3583	1/9833	2/7333	2/3333	2/3509	2/8690
Shahid Beheshti	12	2/3354	1/9833	2/6875	2/6515	2/3246	3/2738
Mashhad	9	2/2583	1/9333	2/5833	2/1919	2/3275	3/3056

Discussion and conclusion

In order to analyze the first research questions "There is a significant difference among the extent of inter- and intra-organization knowledge sharing by faculty members of the universities and research centres affiliated with Ministry of Health Treatment and Medical Education in terms of geographical area where they work", the results indicated no significant difference among faculty members in Tehran and other cities in terms of the extent of intra-organization knowledge sharing. The average extent of sharing knowledge among faculty members in Tehran was higher than that of their counterparts in other cities. However, a significant difference was observed among faculty members in Tehran and other cities in terms of the extent of inter-organization knowledge sharing. Since the average extent of sharing knowledge among faculty members in Tehran was higher than that of their counterparts in other cities, it might be concluded that faculty members in Tehran play a more significant role in inter-organization knowledge sharing than their counterparts in other cities.

The results of this part of the study have not been observed in any other studies yet. Nevertheless, the reason underlying the significant difference in inter-organization knowledge sharing in universities in capital compared to those in other cities might be concentration of many information and communication facilities and technologies in the capital city. In this case, telecommunication might be easier with other universities and research centres in other cities. Moreover, the educational courses provided in the capital city regarding the way of communicating with other people and organizations and the way of using innovative communication and information technologies seems to be more comprehensive and regular. The other possible reason might be holding most workshops and seminars in different areas in Tehran. In addition, it might be due to the fact that library facilities, access to resources, and references and equipment are more in the capital city than those in other cities. All these factors have caused inter-organization knowledge sharing to be different among faculty members in Tehran and other cities since the context and education which people receive play a significant role in individuals' learning and thereby, sharing knowledge.

In response to the second question that to what extent you cooperate with inter- and intra-organization educational and research centres in sharing knowledge and experience, it was found that cooperation with the intra-organization instructors was higher considering both education and research issues which might be due to less geographical distance and its time- and money- saving nature. On the other hand, considering the fact that instructors in an organization more directly and indirectly communicate with each other (through participating in discussion sessions, meeting in different parts of the organizations, etc.), more trust is built among them and more cooperation is expected to be made considering the educational and research issues. This is while cooperation among the instructors in different organizations might not be as more and thereby, the extent of their cooperation might be influenced. The managers of the educational and research organizations are recommended to notice this point and make full use of it in designing the interior space of the organization. In this sense, the interior space can be designed in a way that it brings about the highest extent of interactions among faculty and staff members and thereby, enhances their knowledge sharing. In a similar vein, Knoben & Oerlemans (2006), Torre & Rallet (2005), and Gallaud & Torre (2005) concluded that physical and geographical distance would lead to an increase in face-to-face interaction and implicit knowledge sharing which would increase knowledge creation and creativity. Further, Filippo & Torre (2003) pointed to the significant effect of geographical proximity on conducting joint projects in organizations located in an area. Moreover, the results of a study by van Vuuren (2011) revealed the direct relationship between physical proximity and the extent of interaction and sharing knowledge. For instance, those people working in a floor have more interaction than those working in different floors of a building. Contrarily, those groups in different organizations and cities have much less interaction. However, Gouza (2006)'s findings are at odd with those of previous studies. He stated that taking into account the communications technologies, geographical distance would make no impact on knowledge transfer among different organizations. In this regard, it might be said that despite the prevailing existence of several technologies, people are sometimes reluctant to share knowledge and experience which might stem from their lack of motivation.

References

Broekel, T. & Boschma, R. (2012). Knowledge networks in the Dutch aviation industry: the proximity paradox. *Journal of Economic Geography*, 12(2), 409-433.

Cramton, C.D. (2001). The mutual knowledge problem and its consequences for dispersed collaboration. *Organization Science*, 12(3), 346-371.

Filippi, M. & Torre, A. (2003). Local organizations and institutions. How can geographical proximity be activated by collective projects?. *International Journal of Technology Management*, 26(2-4), 386-400.

Hau, L.N. & Evangelista, F. (2007). Acquiring tacit and explicit marketing knowledge from foreign partners in IJVs. *Journal of Business Research,* 60(11), 1152-1165.

Howells, J.R.L. (2002). Tacit knowledge, innovation and economic geography. *Urban Studies*, 39(5), 871-884.

Gallaud, D. & Torre, A. (2005). Geographical proximity and the diffusion of knowledge: the case of ME's in biotechnolog. In A. Koch (Ed.), *Rethinking Regional Innovation* (127-146). Dordrecht: Kluwer.

Gill, J. & Butler, R.J. (2003). Managing instability in cross-cultural alliances. *Long Range Planning*, 36(6), 543–563.

Gouza, A. (2006). *Key factors of knowledge transfer within university spin- off.* Business Economics Department, Autonomous University of Barcelona, Spain Retrieved June 8, 2013 from: http://webs2002.uab.es/dep-economia-empresa/Journadas/papers/2006/pdf.

Greunz, L. (2003). Geographically and technologically mediated knowledge spillovers between uropean regions. *Annals of Regional Science*, 37(4), 657–680.

Jansen van Vuuren, S. (2011). Inter-organizational knowledge sharing in public sector: the role of social capital and information and communication technology. Unpublished doctoral dissertation, Victoria University of Wellington.

Knoben, J. & Oerlemans, L.A.G. (2006). Proximity and inter-organizational collaboration: A literature review. *International Journal of Management Reviews*, 8(2), 71-89.

Lilleoere, A.M. & Hansen, E.H. (2011). Knowledge-sharing enablers and barriers in pharmaceutical research and development. *Journal of Knowledge Management*, 15(1), 53-70.

MacKinnon, D., Cumbers, A. & Chapman, K. (2002). Learning, innovation and regional development: a critical appraisal of recent debates. *Progress in Human Geography*, 26(3), 293-311.

Meisters, C. & Werker, C. (2004). Physical and organizational proximity in territorial innovation systems: introduction to the special issue. *Journal of Economic Geography*, 4(1), 1-2.

Mitchell, R., Burgess, J. & Waterhouse, J. (2010). Proximity and knowledge sharing in clustered firms. *International Journal of Globalization and Small Business*, 4(1), 5-24.

Nieminen, H.T. (2005). Successful inter-organizational knowledge transfer: developing pre-conditions through the management of the relationship context. *Proceedings of the 21st Annual IMP Conference*, Rotterdam, Netherlands.

Oerlemans, L.A.G., Meeus, M.T.H. & Boekema, F.W.M. (2001). On the spatial embeddedness of innovation networks: an exploration of the proximity effect. *Tijdschrift voor Economische en Sociale Geografie*, 92(1), 60-75.

Rallet A. & Torre A. (2005). Proximity and localization. *Regional Studies*, 39(1), 47-60.

Seyyedeh, N., Daneshgar, F., & Aurum, A. (2009). Investigating inter-organizational knowledge sharing intention in supply chain partnership. *Proceedings of the 20th Australasian Conference on Information Systems*. Australia, Melbourne.

Torre, A. & Gilly, J.P. (2000). On the analytical dimension of proximity dynamics. *Regional Studies*, 34(2), 169-180.

Torre, A. & Rallet, A. (2005). Proximity and localization. *Regional Studies*, 39(1), 47-59.

Van Wijk, R., Jansen, J.J.P., & Lyles, M.A. (2008). Inter-and intra-organizational knowledge transfer: a meta-analytic review and assessment of its antecedents and consequences. *Journal of Management Studies,* 45(4), 830-853.

Janet C.N., Wee, A., & Chua, Y.K. (2013). The peculiarities of knowledge management processes in SMEs: the case of Singapore. *Journal of Knowledge Management*, 17(6), 958-972

Xianyue, L. & Rui, S. (2013). Researching on the Situational Factors Influencing Knowledge Tansfer among Enterprises in Clusters*. Information Technology Journal*, 12(14), 2628-2634.

The Relationship between internet attention and market share of operation systems for personal computers

Li Gu, Weichun Yan and Shule An

WISE Lab, School of Public Administration and Law, Dalian University of Technology,
Dalian, P.R.China, 116024
guli@dlut.edu.cn

Abstract

This paper explores the relationship between network attention and market share of operation systems for personal com-puters using webometrics. Results show that Market share of XP and Win7 network attention, Win7 market share is nega-tively strong correlated, while Win7 market share and the market exhibit a strong positive correlation. Win7 market share and Mac OS market is positive correlated, from which we can infer related problems which users confront often through the Win7 operating system to query the Mac OS. Mac OS operating system market attention and other operating system mar-ket attention have no significant correlation, which show that the Mac OS system are different from other operating sys-tems with relative independence.

Keywords: Operation system; Network attention; Market share; Baidu index

Introduction

Network Attention Can Reflect Social Facts

Baidu Index is an online software based on internet data analysis, developed by the biggist Chinese search engine-Baidu corporation. It is a free data research server based on Baidu webpage search and Baidu news search, which can reflect different key words in user attention and media attention for a spell of time in the past and change tendency to discover share and explore the most valuable news and information on the internet , and reflect social hot topic and interest and need of netizen directly and impersonaly. Baidu search is the biggistChinese network search engine and it's users spread widely around China. Though related data produced by Baidu Index could not be precise beacause of search sampling and approximate calculation, it is true that the tendency came from Baidu Index has scientific basis. Thus, as a search measure to analogy realated problems, Baidu Index is used in a wide field more and more extensively.

Network Attention And Operation System Market Share Have Correlation

Operation system attention and market share of windows XP operation system, Win7 operation system, Vista operation system, Win8 operation system, Mac operation system and Linux, Unix and Netware have various relations. There is positive correlation, as is reflected in the mutual effect of Win7 OS attention and its market share; there is also negative correlation, as is reflected in the mutual effect of XP OS market share and Win7 internet attention. Specific contention will be explained in the fourth chapter.

Literature Review

Literature Review On Operation System

Operation system is a computer procedure to manage and control computer hardware and software, a basic system software to run immediately on bare computer. Any other software could not run unless get the support of operation system. Operation system concentrate on four series: Linux series, Unix series, Netware and windows series. Microsoft Windows is a kind of desktop operation system produced by Microsoft company, and becomes the most popular operation system in the eyes of users from its birth in 1980s. Its subordinate operation system Win7 OS and XP OS have the largest percentage of the operation system market share, higher than 90%, and Vista OS the 0.3%, and Win8 OS quickly takes up 3% of the total share after its birth several years ago. Mac OS is a kind of operation system run in apple series computer based on Unix OS core, developed by apple corporation, compromising 3.5% of the total share. The most prevalent operation system –Win7 OS, XP OS, Vista OS, Win8 OS and Mac OS, have different practical capacity, safety, stability and openness, so they applied in different user groups. Thus, this part aims at the main stream-XP OS, Win7 OS, Vista OS, Win8 OS, Mac OS, Linux and Unix and Netware.

Windows XP OS is a most prevalent operation system, not only dominate absolutely in personal OS, but also powerful in internet OS. XP OS is one of the most excellent OS explored by Microsoft corporation. It has powerful function and superior performance and have shorter starting up time compared to other OS in the same configuration of computers.

Win7 OS receive more and more support from young users because of its gorgeous operation interface and intelligent user interactive and powerful compatibility. Among all its stars developed by Microsoft, the brightistIs that Win7 OS, innovating the user safety system, displacing the former picture identification and identity identification, enhancing the safety of user procedures, simplifying the difficulty of understanding. By virtue of these advantageous characters, Win7 OS is more and more prevalent among IT engineers. Win7's advantage as the new generation OS published by Microsoft in the function view. Because of its stability , safety and ease to use meet the need of most users, displacing Windows XP OS gradually. Wildstrom, Stephen H [1]This article presents information about Microsoft's Windows 7 operating system that is due to be launched on October 22, 2009. The new software includes many improvements and upgrades over both Windows XP and Vista operating systems and is quicker and easier to use. The author has been testing Windows 7 and has found very few compatibility problems. The task bar has also been improved by showing only icons for programs that users are running. Other features are also described.

Windows Vista improve safety defence to make sure the users could have new safety experience, besides its brand new operation experience. Users can resist attack with new approach and maintain the integrity meanwhile, making its confidentiality and usability promoting to a new stage and rendering its users experiencing unprecedented safety. The safety condition of Vista though the mechanism and process to realize safety. Vista OS enhance former edition's safety characters and increase some new characters too, such as user account control, minimum limits of authority strategy, data protection, safe stating up and internet visiting protection. Jones, T [2]thinks Microsoft introduces a major new version of Windows, almost everyone who uses computers, whether for work or play, eventually need to understand how it affect them. The paper outlines how Vista impacts engineers and scientists, to help determine whether you should adopt the new technology. It also describes best practices for developing engineering and scientific applications based on Vista. One of the stated goals of the Windows Vista release is to improve the security of the Windows operating system. Instant search provides advanced tools for designing more specific searches. In Windows Vista, there is a new interface for interacting with the operating system .NET Framework 3.0 (formerly known as WinFX). Now based on Microsoft .NET technology this interface was completely redesigned to be easier to use and more consistent across all Windows Vista features.

We use one word to describe the character of Win8 OS, because it is an operation system with revolutionary change. The successful development of this OS, will let daily computer operation easier and swifter, let users enjoy the happiness of scientific progress and have applicable work environment. Win8 OS give consideration to personal users and corporate users. To personal users, Win8 has faster running rate, easier operating interface and excellent user experience. To corporate users, Microsoft let Win8 OS meet the need of businessmen, and optimize management need in IT field.

Mac OS is a banner OS around the world, based on Unix OS with simple and practical interface. These new characters let Mac OS distinguish from other OS, though existing the shortcoming that it could not support software from other OS. Apple brand computer and its Mac OS is prevalent by virtue of the excellent capacity of stability, expandability and usability. Friedman[3] thinks that The article presents suggestions on optimizing a Mac computer's performance. Explanation on sharing text shortcuts between iOS and OS X devices is given. It is advised to invoke OS X's App Expose feature via keyboard shortcut to view all of the windows for the frontmost application. Suggestion on improving reading in Preview is also given.

Linux is a kind of new type network OS. Its biggist character is open source code, free application programs and biggist advantage the safety and stability. Now it is mainly used in intermediate or advanced network server. Parloff[4] discovers that Today more than 90% of the Fortune 500 rely on Linux in some aspect, according to Jim Totton of Red Hat, the largest vendor of Linux support services. Linux is the "coal and steel of the Information Age," explains Jim Zemlin, executive director of the Linux Foundation, a corporate consortium whose largest contributors include IBM, Intel, Oracle, and Samsung.Linux is an open-source operating system, written by thousands of independent developers working in concert. Users grant one another certain freedoms, like the right to see the source code, alter it, copy it, and redistribute it--all without paying any licensing fees.

Lipschutz, Robert P[5] thinks that NETware provides a great end-user experience. Users of NetWare 6 will notice positive changes in the way they access files and printers compared with previous experiences with NetWare. The resources are now accessible through a browser or directly through Linux, Macintosh, and Windows OSs—with no NetWare client required. That's a maJor improvement for users and a relief for administrators doing installs or upgrades. THIBODEAU[6] thinks that Unix at long last may be on the road to obsolescence, but it's still not clear what will replace it. Gartner reports that its clients are planning to migrate away from Unix. And while some may take two, three or even five years or more to wean themselves off of the venerable operating system, the end is in sight.

Literature Review On The Relationship Between Internet Attention And Market Share

Generally speaking, the more market share a brand has, the more consumers the brand has, the more information on the internet. On the other hand, the more information of a band and its product, the more brand attention and amount of neticizen, the more popularity of the brand. If the praise and criticism of presentations have a proportion of the moderate, that shows popularity and good reputation are appropriate. Thus ,attentionand market share have relationship, but the relationship is not very precise, and could not be expressed by math formulation. The familiar theory can be also be applied in the research of the relationship between the relationship of OS and market share.

MaJunli[7] analysis the relation between passenger flow volume and the change of internet attention by traveler and marginal effect through relevant data, structuring a space-time frame between passenger flow volume and internet attention. The result shows that traveler internet attention has a positive relationship to the change in space and time of passenger flow volume. she also analysis the relationship between passenger flow volume of tourist destination and internet attention by the theory of consumer purchase decision. She describes the change curve of the relationship between daily passenger flow volume during travelling season and internet attention, structure a model in non-structural method to judge the dynamic relation of the two variables. Result shows that there is a bothway connection in the relationship between internet attention and passenger. Li Shixia[8] think that the Internet has become the main channel for most current residents before travel related information. Analysis Baidu index based data, network analysis of tourism in the attention factors change and influence, has certain significance for guiding the development of tourism. In the case of Qingdao, the Qingdao tourism network retrieved 2011 attention curve, found that in 2011 Qingdao tourism network note in time distribution, the key factors affecting the degree of attention to network tourism port -- seasonal changes, leisure time and travel port to have already known, the comprehensive function of the three, affecting the people to travel port more attention.

Network attention and local development related literature review Tension in the background of the rapid development of Internet technology, based on the "Baidu index" is a new means of Web analysis, think of a regional network attention largely reflects the regional image of the region, economic status, creativity, openness and efficiency of social activities, and have a profound the influence on the regional social and economic development, enough to cause more concern and attention. Taking Zhenjiang as an example, the regional network attention in-depth analysis, to explore its causes, and put forward relevant suggestions for improvement.

Investment and network attention degree correlation literature review. LiLongjie[9] believed that each stock investors concerned about the extent, are available through the stock every day is ten times the financial web site Chinese most influential mentioned to measure. The number of new variables mentioned to measure investors limited attention insufficient, can improve the proxy variables selected in the previous researches. Through the empirical analysis of the China two stock markets, using the SVAR econometric model, found the relationship between attention and stock investors have all kinds of connections with the rate of return. Wang Yong studied the relationship between the network attention and stock returns in china. Based on the behavioral finance theory hypothesis, then the network test data. The results show that the high attention of stock in the day and after a day with high volume; the high attention of the rate of return has positive effect, but after one day yield but negatively effect. Conclusions indicate that China's stock market has attention based on the purchase behavior, and such behavior may be institutional investors to gain profit, this provides reference for our investors and regulators in decision-making.

Research methods

Model

This paper uses the SPSS statistical analysis software, with paired samples statistics, paired sample correlation coefficient, the pairwise difference into analysis of difference, paired sample test and Pearson correlation, the operating system to the operating system to browse search volume accounted for more than the amount of processing, finally obtains the linear regression equation.

Data Description

Operating system search volume data is obtained through Baidu index. The main interface in Baidu index, respectively, Win7, VISTA, XP Win8 and Mac OS, and Linux, Unix and Netware, search volume will be 8 operating system, and after the three unified for other operating system. Then the time set for 1 months, and then record the search volume of the month, the final summary.

Baidu statistics is Baidu launched a free professional website traffic analysis tools, can tell the user the visitor is how to find and browse the website. Operating system browsing accounted for more than the data, from Baidu statistics covered more than 1500000 of the site, to help the industry to understand the Internet industry basic data distribution and trend. As is showed in figure 1 to 2

From the operating system market share statistics can be seen in figure, the XP operating system and Win7 operating system is in the dominant position, basically a monopoly of the operating system market. The market share of XP operating system from 2013 June to 2014 May 68% dropped to 57%, while the share more than half but decreases. Win7 operating system rising each year modeistely. The Win8 operating system and Vista operating system and Mac OS operating system and other operating system the proportion of small, and the change is not very obvious.

From the operating system network attention statistics can be seen in figure, Win7 search volume is the highest, followed by Win8, XP operating system and other operating system including Linux, UNIX and Netware, then, Mac OS operating system and Vista operating system has received less attention.

Figure 1. Operation system internet attention

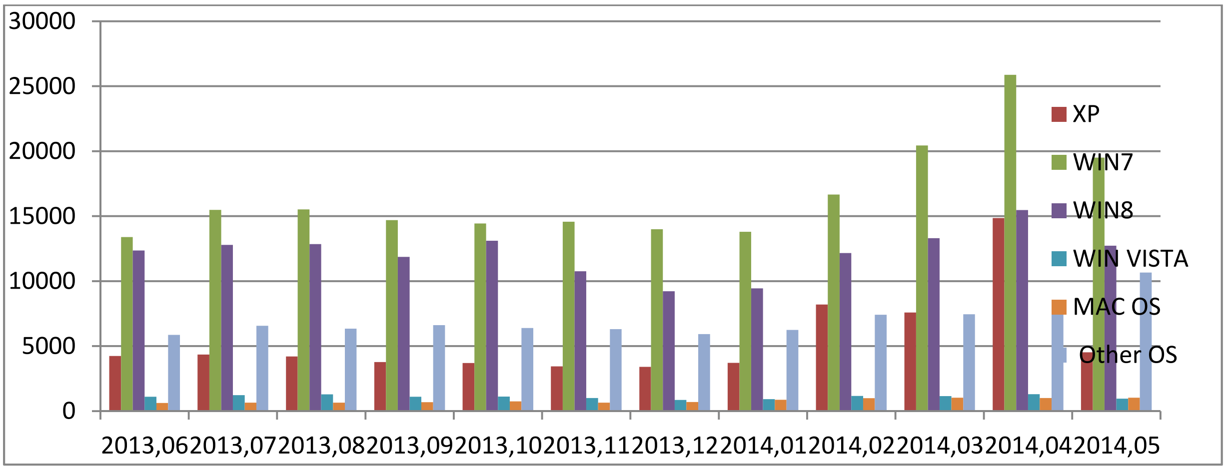
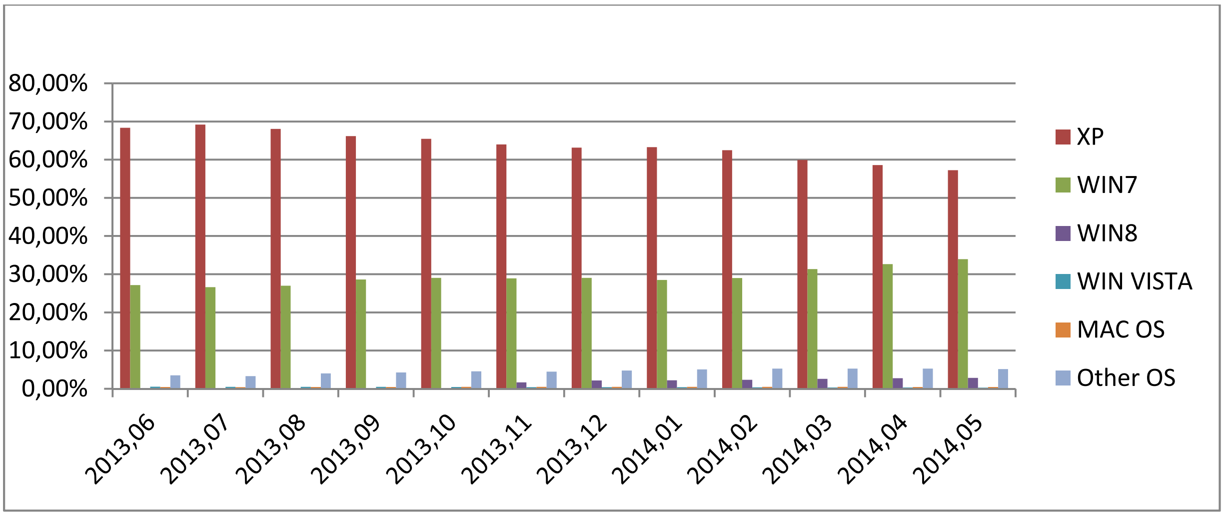


Figure 2. Operation system market share



Data Analysis

Table 3. Correlation Analysis between Page View and Search Range

		XP Page View	WIN7 Page View	WIN8 Page View	WINVISTA Page View	MACOS Page View	Others Page View
XP Page View	Pearson Relation	1	-.954**	-.644*	.972**	-.562	-.894**
	Sig.		.000	.024	.000	.057	.000
WIN7 Page View	Pearson Relation	-.954**	1	.725**	-.901**	.375	.761**
	Sig.	.000		.008	.000	.230	.004
WIN8 Page View	Pearson Relation	-.644*	.725**	1	-.598*	.087	.380
	Sig.	.024	.008		.040	.787	.223
WINVISTA Page View	Pearson Relation	.972**	-.901**	-.598*	1	-.531	-.890**
	Sig.	.000	.000	.040		.076	.000
MACOS Page View	Pearson Relation	-.562	.375	.087	-.531	1	.806**
	Sig.	.057	.230	.787	.076		.002
Others Page View	Pearson Relation	-.894**	.761**	.380	-.890**	.806**	1
	Sig.	.000	.004	.223	.000	.002	
XP Search range	Pearson Relation	-.542	.525	-.008	-.578*	.178	.478
	Sig.	.069	.080	.981	.049	.580	.116
WIN7 Search range	Pearson Relation	-.724**	.772**	.329	-.753**	.067	.542
	Sig.	.008	.003	.297	.005	.837	.069
WIN8 Search range	Pearson Relation	-.197	.370	.085	-.211	-.306	.024
	Sig.	.540	.237	.792	.509	.334	.941
WINVISTA Search range	Pearson Relation	.230	-.131	-.345	.137	-.432	-.244
	Sig.	.473	.684	.271	.670	.161	.444
MACOS Search range	c	-.885**	.819**	.520	-.892**	.550	.854**
	Sig.	.000	.001	.083	.000	.064	.000
Others Search range	Pearson Relation	-.747**	.837**	.911**	-.736**	.132	.499
	Sig.	.005	.001	.000	.006	.682	.098

		XP Search range	WIN7 Search range	WIN8 Search range	WIN VISTA Search range	MAC OS Search range	Others Search range
XP Page View	Pearson Relation	-.542	-.724**	-.197	.230	-.885**	-.747**
	Sig.	.069	.008	.540	.473	.000	.005
WIN7 Page View	Pearson Relation	.525	.772**	.370	-.131	.819**	.837**
	Sig.	.080	.003	.237	.684	.001	.001
WIN8 Page View	Pearson Relation	-.008	.329	.085	-.345	.520	.911**
	Sig.	.981	.297	.792	.271	.083	.000
WIN VISTA Page View	Pearson Relation	-.578*	-.753**	-.211	.137	-.892**	-.736**
	Sig.	.049	.005	.509	.670	.000	.006
MAC OS Page View	Pearson Relation	.178	.067	-.306	-.432	.550	.132
	Sig.	.580	.837	.334	.161	.064	.682
Others Page View	Pearson Relation	.478	.542	.024	-.244	.854**	.499
	Sig.	.116	.069	.941	.444	.000	.098
XP Search range	Pearson Relation	1	.893**	.691*	.555	.648*	.322
	Sig.		.000	.013	.061	.023	.307
WIN7 Search range	Pearson Relation	.893**	1	.738**	.450	.746**	.621*
	Sig.	.000		.006	.142	.005	.031
WIN8 Search range	Pearson Relation	.691*	.738**	1	.824**	.362	.404
	Sig.	.013	.006		.001	.248	.193
WIN VISTA Search range	Pearson Relation	.555	.450	.824**	1	.036	-.019
	Sig.	.061	.142	.001		.910	.954
MAC OS Search range	Pearson Relation	.648*	.746**	.362	.036	1	.729**
	Sig.	.023	.005	.248	.910		.007
Others Search range	Pearson Relation	.322	.621*	.404	-.019	.729**	1
	Sig.	.307	.031	.193	.954	.007	

**. .01 Level (double side) *. 0.05 Level (double side)

Conclusion and Discussions

Market share of XP and Win7 network attention, Win7 market share is negatively strong correlation. From this point of view, Win7 operating system is a replacement of XP operating system and upgrading of products.

Concerned about the Win7 market share and the market presenting a strong positive correlation, that shows the relationship between the market share of Win7 operating system will increase with the increased attention.

Win7 market share and Mac OS market is a strong positive correlation, from which we can infer related problems which users confront often through the Win7 operating system to query the Mac OS.

Mac OS operating system market attention and other operating system market attention have no correlation, show that the Mac OS system with the properties of its own, different from other operating system, with relative independence.

Acknowledgments

This work was supported by the project of "Fundamental Research Funds for the Central Universities" (DUT13RW409) and the project of "Social Science Foundation of Dalian City"(2012dlskyb073).

References

[1] Wildstrom, Stephen H. Win 7: Microsoft Gets It Right (Finally)[J]. *BusinessWeek*. 2009.10:71-72.

[2] Jones,T. Vista vision [Microsoft Windows Vista][J]. *Computing & Control Engineering*. 2007.18 :22-25.

[3] Frideman. Mac OS X Hints[J]. *Macworld*. 2014.31:88-89.

[4] Parloff, Roger. h LINUX CONQUERED THE 500[J]. *Fortune*. 2006.167:82.

[5] Lipschutz, Robert P.Lopez, Dave. Net Ware 6: Robust, but Administration Could Be Easier[J]. *PC Magazine*. 2001.20:44-45.

[6] THIBODEAU, PATRICK.Unix Is Fading, But Its Successor Has Yet to Emerge[J]. *Computerworld*. 2014(48):4.

[7] MaJunli. City of Domestic Tourists and Visitors Network Attention Space-time Correlation Analysis[J]. *Economic geography*. 2011.31:681-685.

[8] LiShixia. Analysis of Influencing Factors on Baidu index of the Tourism Destination[J]. *Journal of Capital Normal University(Natural Socienc;e Edition)*. 2014.13:33-35.

[9] LiLongjie. Stock Returns and Network Media Attention[J]. *Journal of Guangxi University(Philosophy and Social Science)*. 2013.35:56-58.

Impact of articles in non-English language journals – A bibliometric analysis of regional journals of China, Japan, France and Germany in Web of Science

Liu Xiaomin*, Sun Yuan** and He Jing*

*National Science Library, Chinese Academy of Sciences Beijing, China 100190
liuxm@mail.las.ac.cn

**National Institute of Informatics, Tokyo, Japan 101-8430
yuan@nii.ac.jp

Abstract

In the internationalized process of academic communication of scientific community, English has gradually become the main writing language of scientific communication, and the English-language scholarly publication is the main carrier for academic communication. How's the situation of scientific journals in non-English speaking countries? We select the countries of China, Japan, France and Germany and analyze the feature of journals in different languages based on SCI and JCR data. Journals in different languages have different features. From the perspective of the country, the features of China and Japan are more consistent. Journals in different languages of the two countries have the feature of localization. The features of France and Germany are more similar. The feature of localization is gradually strengthened in English-language journals, Multi-language journals and National-language journals. As to journals, usually more local a journal is, weaker its international impact is. However, from the perspective of the fact that journals serve the readers, the publishing model of journals is determined by national input and output of the scientific research. The national scientific achievements need to be communicated in both international and domestic. It is necessary for journals to take the style of diversification.

Introduction

In the internationalized process of academic communication of scientific community, English has gradually become the main writing language of scientific communication, which can be corroborated by the journal language distribution. By Ulrich's website, there are 34,600+ journals under publishing in the field of science all over the world, and English-language journals has 24,000+, which accounts for 69.19% of the total publishing. English has an absolute advantage in the language of academic publications, and the English-language academic publication is the main carrier for academic communication.

Zitt (1998) pointed that the last decade has witnessed the transition from "National Science Model" to the "Transnational model" in scientific communication and publication. During the transition of the mode, the language of scientific communication and publication activity transforms from national language to English. The phenomenon has been verified by the changes of the Chinese scientific journals in last decade. In 2002, there are only 25 English-language journals published in China (Zhang, Wang & Lin, 2003). While in 2013, the number is increased to 252, which is 10 times of ten years ago.

At the time English becomes the mainstream in scientific community, the scientists in non-English speaking countries are puzzled by the language expression. The France scholar Vic Norris (2012) wrote a paper "Scientific Globish: clear enough is good enough" in Trends in Microbiology. In his paper, he mentions that writing in English is a problem to many scientists. And the academic community should transform English to the simplified and standardized English, which is easy to learn, easy to use, and easy to academic communications. With the importance of English communication and the trouble of scientists' native language, academic journals, the carrier for academic papers, are also facing the same problem. The data stated above indicates the dominant status of English-language journals. Well, what is the status of the journals in non-English speaking countries in the international scientific communication? We choose four countries of China, Japan, France and Germany, and analyze the characteristics of journals in different languages published in the four non-English speaking countries.

Data and Overview

The reason for selecting China, Japan, France and Germany is their official language unique. Meanwhile, the amount of journals of these four countries comes on the top of SCI source journals. They have a representation among non-English speaking countries.

The same feature of China, Japan, France and Germany (CJFG) is that they have both National-language journals and English-language journals. And the highest proportion of English-language journals is published by Germany. In terms of the statistics from Ulrich's website, the total number of journals published in Germany is 1844, and English-language journals have 979, accounting for 53.09%. The total number of journals published in France is 659, and English-language journals have 217, accounting for 32.93%. The total number of journals published in Japan is 1619, and English-language journals have 621, accounting for 38.36%. According to Chinese General Administration of Press and Publication, the total number of journals published in China is 4953, and English-language journals have 252, just accounting for 5.14%, which is the lowest. Statistics shows that journal publication in Germany is balanced in languages, the publication language features of France and Japan are very similar. National-language journals in both countries account for approximately 2/3, and national language takes dominant place in China.

It's obvious that SCI is widely used all over the world and journals can be acknowledged by more scientists and are more likely to be cited by more journals if they become source journals of SCI. According to the statistics in JCR 2012, there are 7289 English language journals of 8471 journals, which accounts for 86.04% of the total JCR journals. English has become the main language of scientific communication and English language journals play an essential role of the global core journals. Then what is the status of National-language journals in international communication? We make some further analyses based on JCR data.

The article uses the 2010-2012 data of JCR and Web of Science. During the three years, Chinese journals account for 1.7% of all journals in JCR, and the rates of France, Germany and Japan respectively are 2.3%, 6.7%, 2.9%.The rate of the amount of journals in CJFG account with journals embodied in SCI is the same as the rate of English-language journals in the four countries account with the total native journals publishing.

One of the selection criteria of SCI source journals is that the journals are English full-text journals. "English is the universal language of science. For this reason Thomson Reuters focuses on journals that publish full text in English, or at very least, bibliographic information in English. There are many journals covered in *Web of Science* that publish articles with bibliographic information in English and full text in another language. However, going forward, it is clear that the journals most important to the international research community will publish full text in English. This is especially true in the natural sciences. There are notable exceptions to this rule in the Arts & Humanities and in Social Sciences topics. This is discussed further below. Nonetheless, full text English is highly desirable, especially if the journal intends to serve an international community of researchers. In addition, all journals must have cited references in the Roman alphabet". (The Thomson Reuters Journal Selection Process (http://wokinfo.com/essays/journal-selection-process)

From multiple perspectives, although English journals hold the absolute position, the source journals have the feature of multi-language. During 2012-2013, there are a total of 1190 journals from CJFG in JCR (See Table 1). The feature of multi-language of the SCI source journals is reflected in the language distribution of journals from CJFG (See Table 2). English journals in China, Germany and Japan are in dominant position. Another feature of JCR journals from Germany is that multi-language journals have a high proportion. France has its own outstanding feature, which is that the proportion of English journals, multi-language journals and national-language journals are nearly equal.

Table 1. Number of journals of China, Japan, France and Germany in JCR, 2010-2012

Country	2010	%	2011	%	2012	%
China	130	1.61%	148	1.78%	143	1.69%
France	189	2.34%	192	2.30%	196	2.31%
Germany	545	6.76%	556	6.67%	563	6.65%
Japan	234	2.90%	240	2.88%	241	2.85%
JCR-Total Journals	8061		8336		8471	

Table 2. Language distribution of journals of China, Japan, France and Germany in JCR, 2010-2012

Country	2010			2011			2012		
	Eng	ML	NL	Eng	ML	NL	Eng	ML	NL
China	110	3	17	128	3	17	122	4	17
France	58	69	62	61	68	63	63	69	64
Germany	335	127	83	346	128	82	353	129	81
Japan	195	27	12	200	28	12	205	26	10

Note: Eng: English-language journals; ML: Multi-language journals; NL: National-language journals. The languages are confirmed based on JCR.

Methodology and Results

Journal is an academic communication platform. The universality of source papers represents journals' openness of communication. We can analyze the range of the academic communication with the proportion of the papers by scientists published in the journals of their own languages, which tells the difference among different language journals. By retrieving in the Science Citation Index-Expanded（SCI-E）and counting the native papers in each journal of CJFG during 2010-2012, we can get the average annual percentage of the native papers account in all papers in each journal of each country.

In table 3, China has the highest proportion of native papers published in national-language journals, followed by Japan. The common feature of papers in China and Japan is that the proportion of native papers with total papers is high, which is nearly or over 2/3, no matter what language the journals are in. This shows relatively strong regional feature. The feature of journals in France and Germany is that the proportion of native papers is very low in English-language journals. The proportion of Germany is only 18.22%, which shows the openness of English-language journals is better. The proportion of native papers in native journals of France and Germany is respectively 74.64% and 81.41%. The common feature of CJFG is that national-language journals highlight the characteristic of the country. The multi-language journals of France and Germany have a significantly different performance with China and Japan. The rate of native papers is lower, which represents diversification.

Table 3. Average annual percentage of native papers in CJFG

Country	En	ML	NL
China	72.89%	87.18%	97.78%
France	25.29%	44.59%	73.64%
Germany	18.22%	37.72%	81.41%
Japan	62.09%	71.64%	80.50%

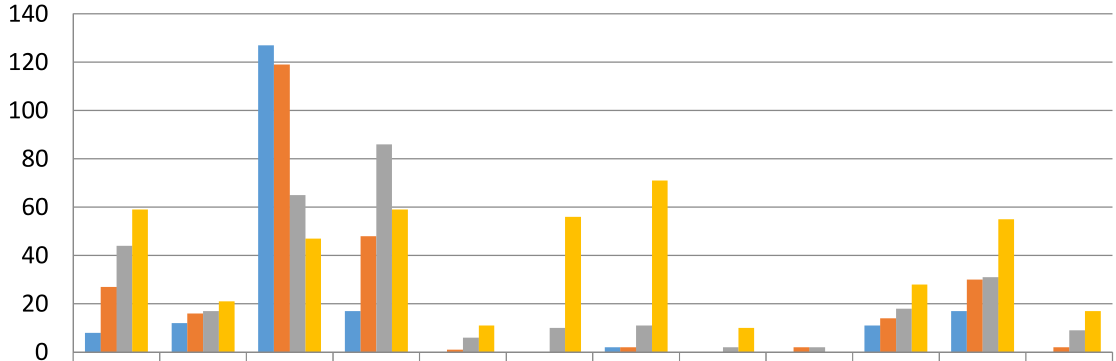
Will the regional feature shows on the journal impact?

We have a further analysis on the international impact of different journals in languages of CJFG using JCR database. From the year 2003, JCR made journal ranking according to the impact factors in the category. The indicator can be viewed a macroscopic and comparable composition for each journal. The bases of the journal ranking are journal's impact factor and subject attribution of the journal. JCR set up 176 categories. In this article, we use the method of JCR journal ranking and analyze the status of different languages journals in the academic communication.

The journals from CJFG are widely distributed in 176 categories, which is too scattered. For ease of comparison, this article astringes the 176 categories to 22 categories according to the mapping table between 22 categories and 176 categories (Thomson Reuters. List of fields for Standard and Deluxe indicators). This is the basis for partition of journals in this article.

The data processing method is stated as follows. First, CJFG's 1190 journals were confirmed their category. Then sequence in descending order according to the three-year average value of impact factor. The amount of journals in each quartile is 1/4 of the total journals of a certain category. Quartile 1 (Q1) has the best performance. By such analogy, Quartile 4 (Q4) contains the tail journals of the category. If a journal is belonging to more than one category, we choose the best performance rank of a category. Analysis the journal ranking, we can conclude the journals' degree of recognition in global scientific communication.

Statistically, the journals published in Germany have a good performance on the whole. The quantity of journals in Q1 is close to the one in Q2. There are 246 journals totally in Q1 and Q2, which accounts for 68.72% of all the German journals. The quantity of journals in Q1 and Q2 exceeds the one in Q3 and Q4. In terms of the proportion of journals in Q 1 with total native journals in each country, the rate of Germany, France, Japan and China is 24.96%, 14.38%, 6.8% and 5%.



	CHI Eng	FRA Eng	GER Eng	JPN Eng	CHI (NL)	FRA (NL)	GER (NL)	JPN (NL)	CHI (ML)	FRA (ML)	GER (ML)	JPN (ML)
Q1	8	12	127	17			2			11	17	
Q2	27	16	119	48	1		2		2	14	30	2
Q3	44	17	65	86	6	10	11	2	2	18	31	9
Q4	59	21	47	59	11	56	71	10		28	55	17

Figure 1. Number of Q1-Q4 journals of CJFG

From the feature of language, the English-language journals distributed in Q1 are better than multi-language journals and national-language journals. Obviously, English-language journals of the four countries are all distributed in Q1. In these four countries, the amount of English-language journals in Germany is in descending order in Q1-Q4 subareas. The amount of English-language journals in France distributed in Q1 is least. But the journals distributed in the four subzones are relatively balanced. The journals of China and Japan distribute rarely in Q1, but the journals distributed in Q2-Q4 subarea are in majority.

In the distribution of multi-language journals subareas, France, Germany and Japan have few or no distribution in Q1, but concentrate in Q4. There are only 4 journals of China distributing in Q2 and Q3. As with the expanding of the subareas, the journals' amounts of each subarea are increasing. The distribution of national-language journals is less unsatisfactory. Only Germany has 2 journals in Q1, which is just 2.32% of all the journals in German. The data shows that national-language journals are mainly concentrated on Q4.

According to the regional distribution of the impact factors, the data of English-language journals are better than the data of multi-language journals and national-language journals for each country. And subareas of multi-language journals have an advantage over the national-language journals to some degree. National-language journals of each country concentrate in Q4 apparently, which indicates the impact of national-language journals has some limitation. The impact is decreasing from English-language journals to multi-language journals then to national-language journals. From the view of data, we have to admit English takes an important place in the scientific communication.

Quartile 1-4 of Impact Factor show the impact degree of journals from one side. We will have a further analysis on the difference in impact of journals in different languages from the component of the impact. The analysis is based on two indicators which are self-citation rate of journals and citation correlation coefficient of native papers.

In recent years, the journal community pays high attention to journals' self-citation. Thomson Reuters (http://wokinfo.com/essays/journal-self-citation-jcr) believes the journal whose self-citation is over 20% is regarded over excessive self-citation according to the statistics on the citation data. In terms of the statistics on journals' self-citation rate offered by JCR, there are a total of 206 journals whose average self-citation is over 20% in 3 years (see Table4) for these four countries, which accounts for 17.32%. In the light of absolute number, Germany has 91 journals whose average self-citation in 3 years is over 20%. On basis of the percentage of journals whose self-citation is over 20% of the total native journals, the highest rate is produced by China which is 29.38%.

Table 4. The journal quantity of each country whose average self-citation rate>20% in 3 years

Journal-Language	China	France	Germany	Japanese
Eng	32	7	29	18
ML	2	8	20	2
NL	13	28	42	5
Sum	47	43	91	25
Ratio	29.38%	21.18%	15.77%	10.00%

Note: Ratio represents the amount of journals whose self-citation rate exceeds 20% accounts of the amount of journals in each country.

The journals' over self-cited phenomenon in Germany and France concentrates on multi-language journals and national-language journals, accounting for 68.13% and 83.72% separately of the total national over self-cited journals. The common feature of China and Japan is that the over self-cited journals are concentrated on national English-language journals.

The self-cited rates of national-language journals from CJFG are outstanding. China has 13 national-language journals whose self-cited rates are too high, accounting for 72.2% of the total national-language journals. The proportion of France, Germany and Japan is 42.42%、48.84% and 41.67% separately. There are two main reasons for the over self-rated phenomenon. On one hand, the human factors may affect. On the other hand, the international influence of the national-language journals is more limited than the English-language journals. The total citation frequency of national-language journals isn't high, so the self-cited rate must be relatively high.

The source papers of academic journals are global. The influence of a journal in academia is affected by the internationalization feature of the component of editors and source papers. There are more than 8000 journals in JCR, publishing the papers of scientific researchers around the world. During 2010-2012, researchers in Germany, France and China published papers in 6200~6800 journals, and researchers in Japan published papers in 5900 journals. The journals publishing papers of these four countries account for 80% of all the journals in JCR. Currently the data basis used to explain the journals' impact indicators come from the cited times contributed by citing journals. How is the influence of the citing journals that publish papers of various countries on the native journals' citation frequency? To illustrate this problem, we conduct the experiment on data statistics.

We analyze the component of the citing journals in CJFG. Although 1190 journals belong to different disciplines, and referencing behaviors are distinguished in different disciplines, there still exists a common phenomenon, which is that few citing journals have high citing frequency. According to this feature, we choose 1190 journals in JCR, and select the citing journals whose citing frequency belongs to TOP20 and select the citation frequency. It should be noted that with the JCR rules of data presentation, when the journal's citing times is lower than a certain threshold, the specific citing journal and its citing frequency should be represented by "All Others". If the citing times of "All Others" belongs to TOP20, it won't be counted in the calculation process in this article. Therefore, some journals only have 19 citing journals in TOP20 actually. Another case is that some journals' citing journals are less than 20. The article ignores the two cases and refers to TOP 20 collectively.

During 2010-2012, 1190 journals have 63,353 records of TOP20. We retrieve the published papers of CJFG in 2008-2012 in SCI-E database, and get the journal list of each country. We match the journals where the papers are published with the TOP 20 journals. And we can get 1190 journals with their citing journals' citing frequency and the amount of papers from China, Japan, France and Germany in the citing journals. In data matching, there are 3636 records of citing journals not in SCI-E database. It is verified that they are the source journals of SSCI or PICI. So there are 59717 valid records actually. The data include the papers whose citation year is previous two years. Namely if the citation year is 2010, the amount of the paper includes papers published in 2008-2009. If the citation year is 2011, the amount of the paper includes papers published in 2009-2010. The data make up the relationship shown in table 5.

Table 5. Example of Relationship between citation frequency of citing journals and amounts of papers

Citing Year	Country	Cited Journal	Citing Journal	Citing times	Amount of papers published previous 2 years
2010	China	J1	Citing_j1	100	20
2010	China	……	…	…	…
2010	China	J1	Citing_j20	200	90
2011	France	J60	Citing_j1	200	30
2011	France	J60	Citing_j20	500	100

Each journal has the data about citation frequency of citing journals and the amount of papers. Using the two matched groups of data, we can calculate the correlation between published papers of a country and the journal's cited times of the country. Correlation coefficient is calculated with the Person formula, and results are shown in Table 6.

Table 6. Correlation coefficient table for paper and its amount in CJFG

Nation	Language	R2010	R2011	R2012
China	Eng	0.64	0.63	0.69
	Multi-Lang	0.67	0.69	0.69
	National-Lang	0.64	0.67	0.70
France	Eng	0.64	0.66	0.66
	Multi-Lang	0.49	0.46	0.43
	National-Lang	0.74	0.76	0.72
Germany	Eng	0.58	0.53	0.59
	Multi-Lang	0.56	0.55	0.57
	National-Lang	0.64	0.59	0.66
Japan	Eng	0.68	0.66	0.66
	Multi-Lang	0.49	0.46	0.43
	National-Lang	0.70	0.61	0.59

From the view of journals in different languages, national-language journals have a strong correlation between cited times and amount of native papers. Native paper means the author's country of a paper on a journal is same as journal publishing country. The correlation coefficient of CJFG is about 0.65. But the correlation coefficient of France is above 0.7 in 3 years, which shows stronger correlation. As to the journals in multi-language journals, the correlation coefficient of China is above 0.67, which shows strong correlation, followed by Germany whose correlation coefficient is above 0.5. The correlation coefficient of France and Japan is lower than 0.5, which shows slightly weaker correlation. English-language journals have a strong correlation between cited times and amount of native papers for each country. China, France and Japan show the consistency, and correlation coefficient of Germany is slightly lower. According to the data, the citation of National-language journals has a greater relationship with native papers, which fits in high percentage of native papers and high self-citation rate of national-language journals and indicates the limitation of native language on international impact.

As to the correlation of the journals in three types of language between citing times and native paper, there are other factors which shouldn't be ignored, such as wide distribution of the papers in JCR journals for each country, inheritance and continuity of the achievements made by native scientific research community, greater international cooperation in scientific research, increase in international collaboration on research papers (The Royal Society, 2011). All these factors strengthen the amount of papers in citing journals during the process of calculating the correlation coefficient. So if there are real cited data of native papers, and the nationality signed on the papers are distinguished, it should be better to explain the extent of the native papers' contribution to the citing behavior, then it will further illustrate the international influence of journals in different languages.

Conclusions and Discussions

Through analyzing journals in different languages in CJFG, we can get the conclusion that journals in different languages have different features. From the perspective of country, the features of China and Japan are more consistent. Journals in different languages of the two countries have the feature of localization. The features of France and Germany are more similar. The localization is gradually strengthened in English-language journals, Multi-language journals and National-language journals.

1. English-language journals have an advantage over the multi-language journals and national-language journals about the international impact for each country. In the quilter distribution of impact factors, English-language journals of each country are distributed in Q1. Excessive self-cited journals are fewer. China and Japan has a higher percentage of native papers in English-language journals, which is over 60%.

2. The multi-language journals of China and Japan are fewer. The feature of multi-language journals in France and Germany is that the amount of native papers is more than the ones of English-language journals. But the proportion is still less than 50%. The impact of multi-language journals is weaker than English-language journals. In the subarea distribution of impact factors, multi-language journals are concentrated in Q3 and Q4. Excessive self-cited journals are increased. Correlation between citation frequency and the amount of native papers is weak. It can be speculated that multi-language journals meet the needs of multi-language readers and authors. The multi-language journal has the transnational and interstate feature, while its readers are limited. The feature of localization is not obvious, but the impact is weaker than English-language journals.

3. The feature of localization is more obviously for national-language journals .The proportion of native papers is nearly or over 80% of the total. Excessively self-cited journals are focused on national-language journals. Citation rate has a strong relationship with the amount of native papers, which fully reflects the limitation of the localization feature of the journals.

The positions in the international scientific communication of journals in different languages tell English-language journals play the leading role in scientific communication. Mintomo Yuasa discovered the transferring law of scientific center in 1962. Since 1920, the world scientific center is transferred from Europe to America. According to the statistics from SCI, the papers from America have been top of the world all along, and English has been the main language of scientific communication. In Zitt and others' research findings, it turns out that the English-language journals' publications are increasing rapidly. Other languages drop down from the main position gradually, becoming more regionalized and localized in academic communities.

The data about journals in different languages from CJFG shows that China and Japan have relatively strong localization feature no matter in distribution of the papers or journals' influence. The localization feature of journals from Germany and France isn't so obvious. On the basis of Wikipedia statistics (http://en.wikipedia.org/wiki/List_of_languages_by_number_of_native_speakers), we can explain this phenomenon with the language distribution. Chinese possesses the largest number of users. But the main distribution areas concentrate on the territory of China. Japanese is limited to the native Japan. Although French is the official language of France, but the users are all through the world, which is the intercontinental language just next to English. Though Germany is focused on the use in Europe, but Germany has the feature of multinational use. Meanwhile, several researchers also believe the scientific center is not unique. England, Germany, France, Japan and other countries are also important geographic distribution of scientific research (Yuan, J.Y., 2005). So the value of scientific research achievement has close relation to language diversity. From the view of language distribution feature, the amount of the languages users isn't the decisive factor. The global multiregional distributions, and the environment of scientific communication are relatively important factors. Multi-language publishing of Germany and France proves. The regional limitation of Chinese and Japanese makes the journals' influence limited in native scientific environment objectively.

The nature of journals is carrying the scientific research achievements, and playing the role of spreading and training. Because the scientific researchers are from different countries, the existence of multi-language journals has its inevitability. If the target of a journal is to improve the international influence, the journal has to be published in English, and the source papers should be global, which forms the international academic exchange platform. If the target is to serve the authors and readers, the native language and multi-language journals are also the publishing modes, which meet the native authors' need, and spread the scientific achievements to wider academic community by exchanging ways in multi-language. China, Japan, France and Germany published Chinese-English, Japanese-English, French-English, French-Spanish, German-English etc. multi-language journals. Maybe this is a good way to improve the international influence of national-language journals or multi-language journals.

With the push of internationalization and globalization, some countries emphasize on citing the papers from "research central countries", rather than the papers from their own countries. For journal publication, some small countries publish their journals only in English (Kirchik, O., Gingras, Y. & Lariviere, V., 2012; Gingras, Y., 2009). But what measurement will the scientific research superpower countries take? The publication model of a country is determined by many factors, such as the input and output for scientific research, rewards from journals and so on. A country's research outcomes sometimes need not only to be communicated internationally, but also to be disseminated internally. Development of journals might need to take the road of diversification.

References

Ulrich's.com, checked in December 16, 2013.

Zitt, Michel; Perrot, Francois; Barre, Remi.The Transition from "National" to "Transnational" Model and Related Measures of Countries' Performance[J]. *Journal of the American Society for Information Science*.1998,49(1):30-42

Zhang, Yuehong; Wang, Sen; Lin, Changdong. (2003) The history of domestic English-language academic journals' development and the implements of internationalization. *Chinese Journal of Scientific and Technical Periodicals*,14 (Special Issue), 754-757.

http://wokinfo.com/essays/journal-selection-process/

Thomson Reuters. List of Fields for Standard and Deluxe Indicators

http://wokinfo.com/essays/journal-self-citation-jcr/

The Royal Society. *Knowledge, Networks and Nations: Global Scientific Collaboration in 21th Century*. 2011

http://en.wikipedia.org/wiki/List_of_languages_by_number_of_native_speakers

Yuan Jianyang. (2005) Review on the Law in Transfer of Scientific Center. *Science Culture Review*, 2(2),1-12.

Olessia Kirchik, Yves Gingras, Vincent Larivière. Changes in Publication Languages and citation Practices and Their Effect on the Scientific Impact of Russian Science (1993–2010)**.** *J. Am. Soc. Inform. Sci. Technol.*, 2012,63(7):1411–1419

Gingras, Y., & Heilbron, J. *The European intellectual space From the formation of nation states to globalization, 19th–21st Century*. 2009, pp. 359–389

Ranking and mappping of universities and research-focused institutions worldwide: The third release of www.excellencemapping.net

Lutz Bornmann*, Moritz Stefaner**, Felix de Moya Anegón*** and Rüdiger Mutz****

*Division for Science and Innovation Studies, Administrative Headquarters of the Max Planck Society, Munich, Germany
bornmann@gv.mpg.de

**Eickedorfer Damm 35, 28865 Lilienthal, Germany
moritz@stefaner.eu

***CSIC, Institute of Public Goods and Policies (IPP), Madrid, Spain
felix.moya@scimago.es

****Professorship for Social Psychology and Research on Higher Education, ETH Zurich, Zurich, Switzerland
mutz@gess.ethz.ch

Abstract

Bornmann, Stefaner, de Moya Anegón, and Mutz (2014a) have introduced a web application (www.excellencemapping.net) which is linked to both academic ranking lists published hitherto (e.g. the Academic Ranking of World Universities) as well as spatial visualization approaches. The web application visualizes institutional performance within specific subject areas as ranking lists and on custom tile-based maps. Scopus data were used which have been collected for the SCImago Institutions Ranking. The second, substantially enhanced version of the web application is described in Bornmann, Stefaner, de Moya Anegón, and Mutz (2014b). In this version, the effect of single covariates (such as the per capita GDP of a country in which an institution is located) on two performance metrics (best paper rate and best journal rate) is examined and visualized. This paper describes the third version.

Introduction

In a list of the most prominent rankings Hazelkorn (2013) names 11 different international rankings. The most important source of data used for the various rankings are abstract and citation databases of peer-reviewed literature (primarily Scopus, which is provided by Elsevier and the Web of Science, WOS, from Thomson Reuters). Publication and citation data is used to make a statement about the productivity and the citation impact of institutions (Bornmann, de Moya Anegón, & Leydesdorff, 2012; Waltman et al., 2012). Recent years have seen a number of different approaches which not only put institutions in a ranking list, but also show their performance on a map (Zhang, Perra, Gonçalves, Ciulla, & Vespignani, 2013). The advantage of this visualization is that regions and countries can be explored and searched for excellent institutions (Frenken, Hardeman, & Hoekman, 2009; van Noorden, 2010).

Bornmann, et al. (2014a) have introduced a web application (www.excellencemapping.net) which is linked to both academic ranking lists published hitherto (e.g. the Academic Ranking of World Universities) as well as spatial visualization approaches. The web application visualizes the scientific performance of institutions (universities or research-focused institutions) within specific subject areas as ranking lists and on custom tile-based maps. The second, substantially enhanced version of the web application is described in Bornmann, et al. (2014b). In this version, the effect of single covariates (such as the per capita GDP of a country in which an institution is located) on two performance metrics (best paper rate and best journal rate) is examined and visualized. A covariate-adjusted ranking and mapping of the institutions was produced in which the single covariates are held constant. This paper describes the third version which is similar to the second.

Methods

The study is based on Scopus data collected for the SCImago Institutions Ranking (http://www.scimagoir.com/). To obtain reliable data in terms of geo-coordinates (Bornmann, Leydesdorff, Walch-Solimena, & Ettl, 2011) and performance metrics (Waltman, et al., 2012), we only consider those institutions that have published at least 500 articles, reviews and conference papers in the period 2007 to 2011 in a certain Scopus subject area in the study.[1] Institutions with fewer than 500 papers in a category are not considered. Furthermore, only subject areas offered at least 50 institutions are included in the web application (Arts and Humanities, for example, is not included). We use this threshold in order to have sufficient institutions for a worldwide comparison. The full counting method was used (Vinkler, 2010) to attribute papers from the Scopus data base to institutions: if an institution appears in the affiliation field of a paper, it is attributed to this institution (with a weight of 1).

Table 1. Number of institutions included in the statistical analyses for 17 different subject areas. The mean best paper rate/best journal rate is the mean best paper rate/best journal rate for the institutions within one subject area.

Subject area	*Number of institutions*	*Mean best paper rate*	*Mean best journal rate*
Agricultural and Biological Science	573	0.15	0.54
Biochemistry, Genetics and Molecular Biology	808	0.14	0.49
Chemical Engineering	181	0.13	0.52
Chemistry	541	0.12	0.57
Computer Science	406	0.14	0.29
Earth and Planetary Sciences	363	0.17	0.62
Engineering	669	0.13	0.34
Environmental Science	267	0.16	0.65
Immunology and Microbiology	239	0.16	0.58
Materials Science	444	0.13	0.47
Mathematics	398	0.14	0.40
Medicine	1309	0.17	0.52
Neuroscience	133	0.17	0.54
Pharmacology, Toxicology and Pharmaceutics	102	0.18	0.60
Physics and Astronomy	702	0.16	0.56
Psychology	81	0.19	0.54
Social Sciences	235	0.17	0.47

[1] The first and second versions of the excellence mapping tool looked at the time period from 2005 to 2009 and from 2006 to 2010. Both versions of the excellence mapping tool can be accessed in the current release.

The performance of the institutions is measured with two indicators. The first indicator, called the best paper rate, shows the proportion of publications from an institution which belong to the 10% most cited publications in their subject area and publication year. The second indicator is the ratio of papers that an institution publishes in the most influential scholarly journals of the world (called the best journal rate). The most influential journals are those which ranked in the first quartile (25%) of their subject areas (journal sets) as ordered by the SCImago Journal Rank SJR indicator. While the best paper rate gives information about the long-term success of an institution's publications, the best journal rate describes an earlier stage in the process, the ability of an institution to publish its research results in reputable journals.

Table 1 shows the number of institutions which are considered as datasets for the 17 subject areas in this study. Out of the 27 available subject areas in Scopus, only those are selected for the study which include at least 50 institutions worldwide. For example, 541 institutions within the subject area of chemistry were included in the analyses. The mean best paper rate for these institutions is .12 (12%) and the mean best journal rate .57 (57%).

The citation impact of the publications from the institutions relates to the period from publication to the beginning of 2014.

The data were analysed by using generalized linear mixed model for binomial data, which takes into account the hierarchical structure of data and properly estimates the standard errors (Bornmann, et al., 2014a; Mutz & Daniel, 2007).

The web application was implemented using modern web technologies and Open Street Map[2] data provided through MapBox[3]. It is based on the javascript frameworks backbone.js[4], jquery[5] and d3.js[6].

Results

To be able to explain the performance differences among the institutions in the regression model, we included the following variables as covariates. The covariates are also utilized to create a covariate-adjusted ranking of the institutions:

(1) Proportion of papers from one institution which were produced in an international collaboration
(2) Corruption perception index
(3) Number of residents in a country
(4) Gross domestic product (GDP) per capita of a country

While the international collaboration covariate relates to individual institutions, all the other covariates apply at the level of individual countries.

Figure 1 shows a screen shot of the web application visualizing the results of the multi-level analyses for 17 subject areas. There is a very short description of the visualization displayed in the upper right section of the web application with a link to click for "More information". The page with the detailed description includes the affiliations of the authors of the web application and a link to this research paper. Under the short description of the web application, the user can choose whether to view the first release of our web application, which relates to the publication period 2005-2009, the second with the publication period 2006-2010, or the third with the publication period 2007-2011. Only in the second release, the user can select two

[2] http://www.openstreetmap.org
[3] https://www.mapbox.com
[4] http://backbonejs.org
[5] http://jquery.org
[6] http://d3js.org

different excellence indicators and several covariates. In the upper right section of the web application, the user can select from 17 subject areas for the visualization. Under the selection window for the subject area, there is another for the covariate (for selecting the corruption perception index, for example).

If the user selects a covariate, the probabilities of (i) publishing in the most influential journals (best journal rate) or (ii) publishing highly cited papers (best paper rate) is displayed adjusted (controlled) for the selected covariate. The results on the performance of institutions can then be interpreted as if the institutions all had the same value (reference point) for the covariate in question. Each covariate was z-transformed over the whole data set (with M=0 and S=1), so that the average probability shows the value in which the covariate in question has the value 0, i.e. exactly equivalent to the median. This allows the results of the model with and without the covariates to be compared.

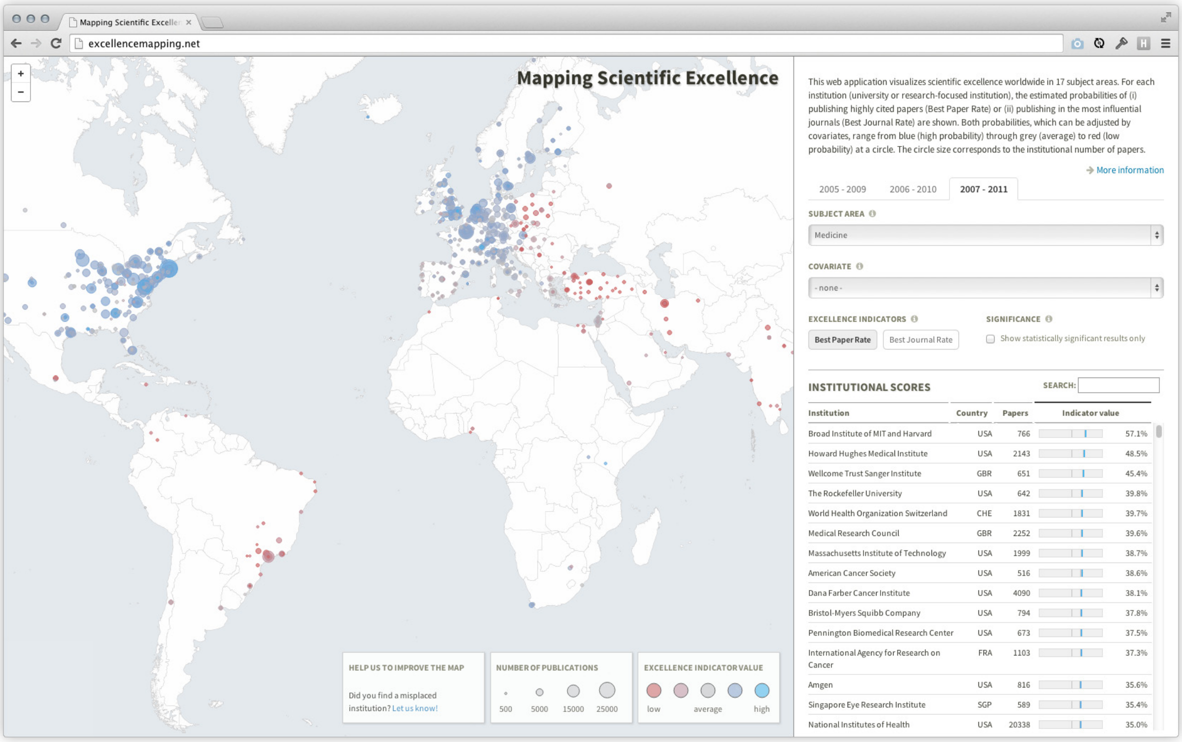


Figure 1. Screen shot of the web application.

Below the selection windows for the subject area and the covariates, users can select one of the two excellence indicators (best paper rate or best journal rate). Our tool shows for each of these indicators the residues from the regression model (random effects) converted to probabilities. In order to have values on the original scale for both indicators for the tool (i.e. proportion of papers in the excellent range or published in the best journals), the intercept was added to the residues. Users can tick "Show statistically significant results only" to reduce the set of visualized institutions in a field to only those which differ statistically significantly in their performance from the mean value.

The map on the left-hand side of the screen shows a circle for each institution with a paper output greater than or equal to 500 for a selected subject area (e.g. Physics and Astronomy). Users can move the map to different regions with the mouse (click and drag) and zoom in (or out) with the mouse wheel. Country and city labels and map details appear only at zoom levels of a certain depth, primarily in order to facilitate perception of the data markers. Zooming can also be done with the control buttons at the top left of the screen. The circle area for each institution on the map is proportional to the number of published papers in the respective subject area. For example, the Centre National de la Recherche Scientifique (CNRS) has the largest circle (in Europe) on the Physics and Astronomy map, highlighting the high output of papers in this subject area. As several circles overlap on larger cities, users can select all the circles in a certain region with the mouse, by holding down the shift key and marking out the area on the map in which the institutions in question are located. These institutions are then displayed on the right-hand side of the web application under "Your selection". The color of the circles on the map indicates the excellence indicator value for the respective institution using a diverging color scale, from blue through grey to red (without any reference to statistical testing): If the excellence indicator value for an institution is greater than the mean (expected) value across all institutions, its circle has a blue tint. Circles with red colors mark institutions with excellence indicator values lower than the mean. Grey circles indicate a value close to the expected value.

All those institutions which are taken into account in the multi-level model for a subject area (section "Institutional scores") are listed on the right-hand side of the web application. The name, the country, and the number of all the papers published ("Papers") are displayed for each institution. In addition, the probabilities of (i) publishing in the most influential journals (best journal rate) or (ii) publishing highly cited papers (best paper rate) are visualized ("Indicator value"). The greater the confidence interval of the probability, the more unreliable for an institution it is. If the confidence interval does not overlap with the mean proportion across all institutions (the mean is visualized by the short line in the middle of "Indicator value"), this institution has published a statistically significantly higher (or lower) best paper or best journal rate than the average across all the institutions ($\alpha = 0.165$). The institutions in the list can be sorted (in descending or ascending order in the case of numbers) by clicking on the relevant heading. Thus, the top or worst performers in a field can be identified by clicking on "Indicator value." Clicking on "Papers" puts the institutions with high productivity in terms of paper numbers at the top of the list (or at the end). In Biochemistry, Genetics and Molecular Biology, for example, the institution with the highest productivity between 2007 and 2011 is the CNRS; in terms of the best paper rate, the best-performing institution is the Broad Institute of MIT and Harvard. The column farthest on the right ("Δ rank") in the "Institutional Scores" section shows for each institution by how many rank places it goes up (green, arrow pointing upwards) or goes down (red, arrow pointing downwards), if the user selects a certain covariate. For example, the Institute for High Energy Physics (RUS) improves its position by 5 places compared to the ranking which does not take the covariate "corruption perception index" into account in the Physics and Astronomy subject area. The ranking differences in this column always relate to all the institutions included. The differences do not therefore change if one looks at only the statistically significant results.

If a covariate has been chosen, one can, for example, sort the institutions by Δ rank, which puts the institutions which benefit most from the covariate being taken into account at the top of the list. Using the search field at the top right, the user can find a specific institution in the list. To identify the institutions for a specific country, click on "Country". Then the institutions are first sorted by country and second by the indicator value (in ascending or descending order). "Your selection" is intended to be the section for the user to compare institutions of interest directly. If the confidence intervals of two institutions under "Indicator value" do not overlap, they differ statistically significantly on the 5% level in the best paper or best journal rate. For example, in Physics and Astronomy, Stanford University and the Helmholtz Gemeinschaft are visualized without overlap (publication years 2007 to 2011). The selected institutions in "Your selection" can be sorted by each heading in different orders. These institutions are also marked on the map with a black border. Thus, both institutional lists and institutional maps are linked by the section "Your selection". For the comparison of different institutions, it is not only possible to select them in the list but also on the map with a mouse click. A new comparison of institutions can be started by clicking on “Clear”.

If the user has selected some institutions or has sorted them in a certain order, the selection and sort order are retained if the subject area is changed. This feature makes it possible to compare the results for certain institutions across different subject areas directly.

Discussion

Compared to the mapping and ranking approaches introduced hitherto, our underlying statistics (multi-level models) are analytically oriented – following Bornmann and Leydesdorff (2011) – by allowing (1) the estimation of statistically more appropriate values for the best paper and best journal rates than the observed values; (2) the calculation of confidence intervals as reliability measures for the institutional performance; (3) the comparison of a single institutions with an “average” institution in a subject area and (4) the direct comparison of at least two institutions. (5) Furthermore, taking covariates into account when mapping and ranking institutions allows an adjusted view of institutional research performance. For example, with our application it is possible to look at the performance of institutions worldwide in countries with the same financial background (that is, the same GDP). This highlights institutions showing a relatively high performance despite a bad financial situation in the country (Bornmann, in press).

References

Bornmann, L. (in press). Ranking institutions by the handicap principle. *Scientometrics*.

Bornmann, L., de Moya Anegón, F., & Leydesdorff, L. (2012). The new Excellence Indicator in the World Report of the SCImago Institutions Rankings 2011. *Journal of Informetrics, 6*(2), 333-335. doi: 10.1016/j.joi.2011.11.006.

Bornmann, L., & Leydesdorff, L. (2011). Which cities produce more excellent papers than can be expected? A new mapping approach—using Google Maps—based on statistical significance testing. *Journal of the American Society of Information Science and Technology, 62*(10), 1954-1962.

Bornmann, L., Leydesdorff, L., Walch-Solimena, C., & Ettl, C. (2011). Mapping excellence in the geography of science: an approach based on Scopus data. *Journal of Informetrics, 5*(4), 537-546.

Bornmann, L., Stefaner, M., de Moya Anegón, F., & Mutz, R. (2014a). Ranking and mapping of universities and research-focused institutions worldwide based on highly-cited papers: A visualization of results from multi-level models. *Online Information Review, 38*(1), 43-58.

Bornmann, L., Stefaner, M., de Moya Anegón, F., & Mutz, R. (2014b). What is the effect of country-specific characteristics on the research performance of scientific institutions? Using multi-level statistical models to rank and map universities and research-focused institutions worldwide. Retrieved January 16, 2014, from http://arxiv.org/abs/1401.2866

Frenken, K., Hardeman, S., & Hoekman, J. (2009). Spatial scientometrics: towards a cumulative research program. *Journal of Informetrics, 3*(3), 222-232. doi: 10.1016/j.joi.2009.03.005.

Hazelkorn, E. (2013). Reflections on a decade of global rankings: what we’ve learned and outstanding issues. *Beiträge zur Hochschulforschung, 35*(2), 8-33.

Mutz, R., & Daniel, H. D. (2007). Development of a ranking procedure by mixed Rasch model and multilevel analysis - psychology as an example. *Diagnostica, 53*(1), 3-16. doi: Doi 10.1026/0012-1924.53.1.3.

van Noorden, R. (2010). Cities: building the best cities for science. Which urban regions produce the best research - and can their success be replicated? *Nature, 467*, 906-908. doi: 10.1038/467906a.

Vinkler, P. (2010). *The evaluation of research by scientometric indicators*. Oxford, UK: Chandos Publishing.

Waltman, L., Calero-Medina, C., Kosten, J., Noyons, E. C. M., Tijssen, R. J. W., van Eck, N. J., . . . Wouters, P. (2012). The Leiden Ranking 2011/2012: data collection, indicators, and interpretation. *Journal of the American Society for Information Science and Technology, 63*(12), 2419-2432.

Zhang, Q., Perra, N., Gonçalves, B., Ciulla, F., & Vespignani, A. (2013). Characterizing scientific production and consumption in Physics. *Sci. Rep., 3*. doi: 10.1038/srep01640.

Infometrics analysis of Scientific-literature in Pediatrics obesity

M.H. Biglu* and M. A-Farhangi**

*Management and Medical Informatics School, Tabriz University of Medical Sciences, Tabriz, Iran
mh_biglu@yahoo.com

**School of Nutrition, Tabriz University of Medical Sciences, Tabriz, Iran
abbasalizad_m@yahoo.com

Abstract

Introduction: Childhood obesity is a serious health issue that has both immediate and long-term effects on health and well-being. Although obesity affects both children and adults, but childhood obesity is more serious than obesity in adults. The objective of current study was to analyse all papers indexed in MEDLINE under the major topic of pediatric obesity.

Methodology: An Informtrics analysis was conducted to plot the development of scientific activities in the field of Pediatrics obesity. Database of MEDLINE was used to extract all papers in the field of pediatric Obesity. Extraction of papers was restricted into major topics "pediatric obesity" from the Search Builder pull-down menu in the advanced search screen, this cause to obtain the articles that their major topics are in the desired subject area.

Results: Analysis of data showed that a total number of 258 papers indexed as major topic of pediatric obesity in MEDLINE. The pediatric obesity as a major topic of Medical Subject Heading was initially introduced in 2013, that is why we are only able to obtain the papers under the major topic of pediatric obesity since this year in MEDLINE. The results of study specified the investigation of pediatric obesity in five major categories: Nutritional-aspect (49%), social-aspect (29%), Epidemiological-aspect (13%) psychological-aspect (5%), and Genetics-aspect (5%).

Conclusion: Although the pediatric obesity was a very new added major topic into Medical Subject Headings (MeSH) in MEDLINE, but it has attracted the consideration of many scientists from all over the world, special from North America and Western Europe. More than 66% of researches in the field were supported by Non-U.S. Governmental organizations. Majority of investigation in paediatric obesity was done in the subject area of nutritional principles. The social field positioned on the second stage of ranking. The genetics and psychology aspects of study seem to be in the inferior concern of scientists. Regarding the important influence of socio-psycho-genetic aspects on the paediatric obesity, these issues of study should be taken more under consideration by policy-makers and nutritional scientists..

Keywords: Infometrics, Pediatric Obesity, MEDLINE

Introduction

Obesity is a global public health problem, associated with a number of chronic disease including cardiovascular disease, metabolic syndrome, type 2 diabetes mellitus and several form of cancers (Caterson & Gill, 2002; Schwarzenberg & Sinaiko, 2006). The prevalence of obesity is increasing worldwide in both developed and developing countries (Flynn et al., 2006). Childhood obesity is another serious health issue that has both immediate and long-term effects on health and well-being; although obesity affects both children and adults (Dietz, 1994), but childhood obesity is more serious than obesity in adults. Obesity in childhood is of particular concern due to its associated health consequences and its influence on young psychosocial development (Must & Strauss, 1999; Power, Lake, & Cole, 1997). Overweight children are more likely to develop obesity in adulthood and approximately 50% of overweight adolescents and over one-third of overweight children remain obese in later ages (Power et al., 1997; Serdula et al., 1993). Currently it was estimated that the prevalence of childhood overweight and obesity range from 12% to over 30% in developed countries and from 2% to 12% in developing countries (Lobstein, Baur, & Uauy, 2004).There are no evidences to reveal the attitude of scientists towards dealing with pediatric obesity. Since the number of published papers in each subject area can reveal the attitudes and attempts of individuals and/or organization towards the same field (Biglu, Ghavami, & Biglu, 2014) ,hence in this study we aim to extract and analyze all papers published in the journals that indexed in the database of MEDLINE under the major topics of "Pediatrics obesity

Methodology

An Infomtrics study was conducted to analyse all papers distributed by the journals that were indexed in MEDLINE." MEDLINE is the National Library of Medicine (NLM) journal citation database. Started in the 1960s, it now provides over 21 million references to biomedical and life sciences journal articles back to 1946. MEDLINE includes citations from over 5,600 scholarly journals published around the world (NLM, 2014)". On 20th April 2014 we extracted all related papers from MEDLINE. Extraction of papers from MEDLINE was restricted into major topics of "pediatric obesity" from the Search Builder pull-down menu in the advanced search screen. This kind of data extracting facilities the way to obtain the articles in the most related subject area (Biglu, 2008).The origin country of papers identified from the field of addresses (AD) in each records, then the articles were geographically classified. All obtained papers went under content analysis by specialists to determine the sub-categories of papers in the field.

Findings

All articles indexed under the major topics of “*Pediatrics obesity*” in MEDLINE were extracted and went under analysis. The study showed that pediatric obesity was initiated as a MeSH term first in 2013; hence the extraction of papers under the major topic of "*pediatric obesity*" is only possible afterward 2013 in MEDLINE. The extraction of data led to 258 articles in the form of Journal Article (246 papers), Research Support (183 papers), Review (35 papers), Comparative Study (23 papers), English Abstract (22 papers), Multicentre Study (15 papers), Randomized Controlled Trial (8 paper), Clinical Trial (7 paper), Editorial (7 paper), Letter 4 paper), Case Reports (3 paper), Controlled Clinical Trial (3 paper),Validation Studies (3 papers), Evaluation Studies (2 papers), Introductory Journal Article (2 papers), Meta-Analysis (2 papers), Comment (1 paper), Interview (1 paper), Practice Guideline (1 paper), and Pragmatic Clinical Trial (1 paper). From a total number of 258 papers the majority of them (91%) was in English language. Only 5% of papers was in German. Some papers found in Spanish (5 papers), French (3 papers), and only paper was in Portuguese (Fig. 1).

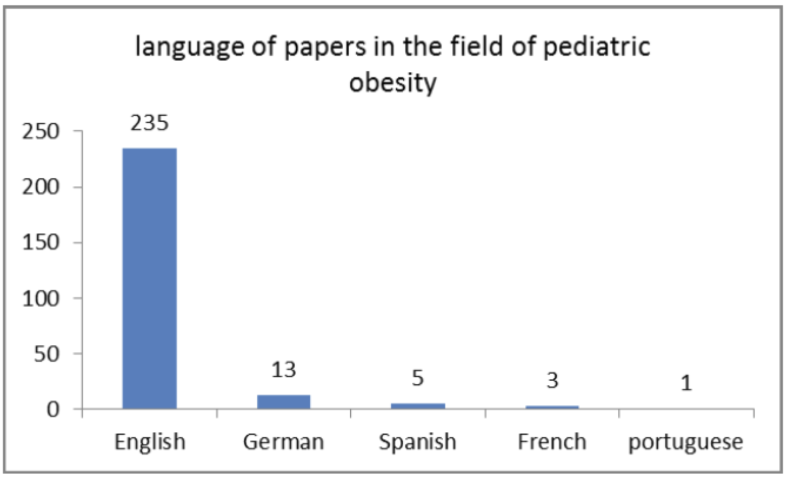


Figure 1. Language of papers in the field of paediatric obesity

A total number of 258 journals contributed papers in the field of pediatric obesity through the period of study

Table 1. Publication types of papers in the field of pediatric obesity.

Publication Type	Number	percent
Journal Article	246	43%
Research Support	183	32%
Review	35	6%
Comparative Study	23	4%
English Abstract	22	4%
Multicenter Study	15	3%
Randomized Controlled Trial	8	1%
Clinical Trial	7	1%
Editorial	7	1%
Letter	4	1%
Case Reports	3	1%
Controlled Clinical Trial	3	1%
Validation Studies	3	1%
Evaluation Studies	2	0%
Introductory Journal Article	2	0%
Meta-Analysis	2	0%
Comment	1	0%
Interview	1	0%
Practice Guideline	1	0%
Pragmatic Clinical Trial	1	0%
Total	569	100%

As shown in table 2, from a total number of 112 journals contributed papers in the major topics of pediatric obesity in MEDLINE, the journal of Childhood obesity (Print) sharing 9% (24 papers) of global publication was the most prolific journal, followed by Journal of obesity 8% (20 papers), Obesity (Silver Spring, Md.) 5% (14 papers) and Bundesgesundheitsblatt, Gesundheitsforschung, Gesundheitsschutz 5% (12 papers).

Table 2. Ten top Journals distributing articles in the major topic of "pediatric obesity" in MEDLINE Rank

Rank	Journal name	Origin Country of Publication	Frequency
1	Childhood obesity (Print)	USA	24
2	Journal of obesity	USA	20
3	Obesity (Silver Spring, Md.)	USA	14
4	Bundesgesundheitsblatt, Gesundheitsforschung, Gesundheitsschutz	Germany	12
5	International journal of obesity (2005)	England	8
6	Journal of health care for the poor and underserved	USA	7
7	Obesity facts	Switzerland	7
8	Nutricionhospitalaria	Spain	6
9	Pediatric obesity	England	6
10	Social science & medicine (1982)	England	6

Authors from 34 countries shared their papers in the major topic of pediatric obesity in MEDLINE. Among them, the American authors were the most productive, sharing 46% of global publication in the major topic of pediatric obesity in MEDLINE. The following authors were from Germany sharing 11% of global publication in MEDLINE. Regarding to the origin regions of published papers, they mostly came from North America (50%) and Western Europe (48%). Table 3 shows the origin country of authors, who shared their works in MEDLINE. The table is restricted into the origin country of 20 top prolific authors in the field.

Table 3. Origin country of authors contributing papers in the field of pediatric obesity

Rank	Country	Paper	Percent
1	USA	123	46%
2	Germany	29	11%
3	UK	16	6%
4	Canada	12	4%
5	Italy	10	4%
6	France	6	2%
7	Spain	6	2%
8	Brazil	6	2%
9	China	6	2%
10	Netherlands	5	2%
11	Romania	5	2%
12	Iran	5	2%
13	Malaysia	4	1%
14	Australia	4	1%
15	Portugal	3	1%
16	Greece	3	1%
17	Chile	3	1%
18	India	3	1%
19	Denmark	2	1%
20	Norway	2	1%

The content analysis of extracted papers in the major topic of pediatric obesity indicated that the scientists' approaches were in 5 major categories. Research in Nutritional aspect was the most frequented subject area, 49% of total research were in this area. 29% of researches was in Social aspects, 13% in Epidemiological aspects, 5% in psychological aspects, and only 4% was in genetically aspects. Some studies were categorized in more than one subject area, that is why the number of papers (283 papers) shown in the table 4 is greater than those extracted from the database of MEDLINE (258 papers).

Table 4. Different kinds of researches in the Major topic of pediatric obesity

Research kinds	Paper	Percent
Nutritional aspect	139	49%
Social aspect	83	29%
Epidemiological aspect	37	13%
Psychological aspect	14	5%
Genetics aspect	10	4%
Total	283	100%

Discussion and conclusion

Analysis of extracted data indicated that the "*pediatric obesity*" first was introduced as a major topic of medical subject heading in MEDLINE in 2013. From a total number of 258 papers indexed as a major topic of pediatric obesity in database of MEDLINE, the majority of them were in the form of journal articles (43%) and Research Support (32%). A great number of research Supports was carried out by Non-U.S. Gov't. This is a hint that non-governmental sections pay more attention on child-hood obesity. 91% of extracted papers were in English language. This pheromone should not come as a surprise, MEDLINE has concerned indexing of papers in English language since many years ago (Biglu, 2007).The American authors sharing 46% of global publication in the database of MEDLINE represented the most dominant authors whose works published in the subject area, followed by authors from Germany distributing 11% of global publication. Regarding to the origin regions of published papers, they mostly came from North America (50%) and Western Europe (48%). One may interprets this occurrence in such a way that developed countries have more facilities for doing research in different areas e.g. in the field of obesity. We should bear in mind that the obesity has a strong relationship with the income level of countries. Base on the recent report of WHO, the prevalence of overweight in high income and upper middle income countries was more than double that of low and lower middle income countries (WHO, 2014).

The content analysis of obtained papers identified that the scientists' methodologies towards pediatric obesity were in five categories. The Nutritional category was the most frequented of them followed by Social category. The Genetics category was the least one. The genetics and psychology aspects of study seem to be in the inferior concern for nutritional-scientists. Regarding the important influence of socio-psycho-genetic aspects on the pediatric obesity, these issues are crucial issues related to the obesity; hence it is strongly recommended the nutritional scientists and policy makers in research centres to take under consideration these issues of study.

References

Biglu, M. H. (2007). The editorial policy of languages is being changed in Medline. ACIMED;16(3).

Biglu, M. H. (2008). *Scientometric study of patent literature in MEDLINE & SCI* (Doctoral), Humboldt-Universität zu Berlin.

Biglu, M. H., Ghavami, M., & Biglu, S. (2014). Authorship, institutional and citation metrics for publications on postmenopausal osteoporosis. *Osteoporos Int, 25*(4), 1337-1343. doi: 10.1007/s00198-013-2603-3.

Caterson, I. D., & Gill, T. P. (2002). Obesity: epidemiology and possible prevention. *Best Pract Res Clin Endocrinol Metab, 16*(4), 595-610.

Dietz, W. H. (1994). Critical periods in childhood for the development of obesity. *Am J Clin Nutr, 59*, 955-959.

Flynn, M. A., McNeil, D. A., Maloff, B., Mutasingwa, D., Wu, M., Ford, C., & Tough, S. C. (2006). Reducing obesity and related chronic disease risk in children and youth: a synthesis of evidence with 'best practice' recommendations. *Obes Rev, 7 Suppl 1*, 7-66. doi: 10.1111/j.1467-789X.2006.00242.x.

Lobstein, T., Baur, L., & Uauy, R. (2004). Obesity in children and young people: a crisis in public health. *Obes Rev, 5 Suppl 1*, 4-104. doi: 10.1111/j.1467-789X.2004.00133.x.

Must, A., & Strauss, R. S. (1999). Risks and consequences of childhood and adolescent obesity. *Int J Obes Relat Metab Disord, 23 Suppl 2*, S2-11.

NLM. (2014). from http://www.nlm.nih.gov.

Power, C., Lake, J. K., & Cole, T. J. (1997). Measurement and long-term health risks of child and adolescent fatness. *Int J Obes Relat Metab Disord, 21*(7), 507-526.

Schwarzenberg, S. J., & Sinaiko, A. R. (2006). Obesity and inflammation in children. *Paediatr Respir Rev, 7*(4), 239-246. doi: 10.1016/j.prrv.2006.08.002.

Serdula, M. K., Ivery, D., Coates, R. J., Freedman, D. S., Williamson, D. F., & Byers, T. (1993). Do obese children become obese adults? A review of the literature. *Prev Med, 22*(2), 167-177. doi: 10.1006/pmed.1993.1014.

WHO. (2014). Obesity:Situation and trends. from http://www.who.int/gho/ncd/risk_factors/obesity_text/en/

Transactions Reports Analysis Islamic Azad University Marvdasht – branch website: A Case Study

Marzieh Yari Zanganeh* and Nadjla Hariri**

*Department of Knowledge and Information Science, Marvdasht Branch, Islamic Azad University, Marvdasht, Iran
myzanganeh@yahoo.com

**Associate professor, department of Knowledge and Information Science, Science and Research Branch, Islamic Azad University, Tehran, Iran
nadjlahariri@gmail.com

Abstract

The main focus Log in user behavior and behavior helps to know what happened. To this end ,the aim of this study is to achieve user behavior by using data mining technology, data were collected during six months in the form of log file in the Islamic Azad University Marvdasht branch website. Research population included pages of the website visitors during 6 months (from February 2013 to July, 2013). Researc h method is transaction reports that were saved on the site server. University website to get the required data from the server and then use the Excel software for filtering unrelated data were analyzed. Results showed that more pages that users have visited is concerned to visit the home page and the second page Hot Hits is educational Chart system. Users using three methods to retrieve information needs. Most visitors use of direct references to the homepage. 27 Percentage of hits has been via search engines.the top used search engine Google and Yahoo the least by visitors. 6 percent of Visitors access to the university website through the references of other Web pages .the most visited hours at 22 pm and 13 pm has the fewest hits.the highest Site visit on Saturday, Tuesday and Wednesday have been done. Firefox Browsers is most used by visitors. Most widely used operating system is Windows 7.

Keywords: Transactions Reports, Islamic Azad University, Marvdasht, Web log file.

Introduction

Though user needs may be elicited in other ways, usage mining of Web logs is a widely used alternative for understanding usage patterns. To this end, search engines collect query logs and toolbar data, while content providers log information about search referrals, site search and browsing activity. In both cases, the information is typically aggregated into sessions, which group together the actions performed by the same user (typically identified using IP addresses or cookies) within a predefined window of time.(Hollink, et al, 2013).

Transaction-log analysis is a valuable and often used method to study user's search behavior on the Web. It is popular for being a non-intrusive way to study large amounts of usage data (Bernard, 2006). On the other hand, it does not provide information about the underlying information need of the searchers (Ronald et al, 1983) or how satisfied they were with the result (Martin, 2002). In the last decade, several approaches have emerged that use semantics to aid the analysis of Web logs.Web mining is the integration of information gathered by traditional data mining methodologies and techniques with information gathered over the World Wide Web (Kosala, et al, 2000). It is used to understand customer behavior, evaluate the effectiveness of a particular Web site, and help quantify the success of a marketing campaign. It also allows looking for patterns in data through content mining, structure mining, and usage mining (Lizhen et al, 2002).Current software application often produce (or can be con_gured to produce) some auxiliary text files known as log files. Such files are used during various stages of software development, mainly for debugging and profiling purposes.

Use of log files helps testing by making debugging easier. It allows to follow the logic of the program, at high level, without having to run it in debug mode. (Andrews, ?)

Problem statement

Web- based technology , the provision of useful features including availability of resources , simply expanded and updated , and keep them on the web , as a suitable technology has been introduced and developed many educational environments around the web world are using it . Although intelligent tools for understanding online user behaviors in order to increase sales and profits have been developed, but little work on the discovery and access to education has been conducted to understand the behavior patterns of online users. Data mining in recent years due to the availability of massive amounts of data, much interest in the scientific community and information industry, has attracted as one of the most recent advances in data management technologies are considered along. With the increasing popularity of the World Wide Web, the bulk of the data collected by web servers to web log file format .The file in which all activities are occurring in the web server, can be as a very rich sources of information to understand and recognize the behavior of Web users, can be used. The main focus Log in user behavior and behavior helps to know what happened.

Importance of research

The web world is unthinkable is expanding every year by the business that is done by the internet, billions of dollars are exchanged. This has led to an important business is to improve the user. Today, the site owners intend to use the personalized environment for users according to their behavior, to attract users. Exploring Web application, as One of the applications data mining techniques in order to improve the design of web sites log files.Web server log files, potentially containing data , experimental websites are useful for improving the efficiency and benefits for some applications, especially in trade, the pick . By analyzing these files can be expected to increase the performance of Web links, which have a positive impact and are very useful for web designers in order to increase user satisfaction to help.

Aims of the research

The aim of this study is to achieve user behavior by using data mining technology, data collected during six months in the form of log file in the of Marvdasht Azad University website.

Method

This survey is a descriptive one with a statistical procedure. The research method is transaction log analysis. The required information was obtained from data and log files of Islamic Azad University Marvdasht branch website server, from February 2013 to July, 2013.Research-community includes pages have visited by the visitor of Islamic Azad University Marvdasht website which are available at the log file server During the 6 month. In this research, first filtering unrelated data, then 500Data for the analysis were selected from among 5000000 data.

Web usage mining has several applications (Naresh, 2003) and is used in the following areas:

- It offers users the ability to analyze massive volume of click stream or click flow data,integrate the data seamlessly, with translation and demographic data from offline sources.
- Personalization for a user can be achieved by keeping track of previously accessed pages. These pages can be used to identify the typical browsing behavior of a user and subsequently to predict desired pages.
- By determining access behavior of users, needed links can be identified to improve the overall performance of future accesses.
- Web usage patterns are used to gather business intelligence to improve customer attraction, customer retention, sales, marketing, and advertisements cross sales.

- Web usage mining is used in e-Learning, e-Business, e-Commerce, e-Newspapers, e-Government and Digital Libraries. The information gathered through Web mining is evaluated by using traditional data mining parameters such as clustering and classification, association, and examination of sequential patterns.

A Web log file records activity information when a Web user submits a request to a Web Server. The main source of raw data is the web access log which we shall refer to as log file. As log files are originally meant for debugging purposes. (Jaideep et. al,2000)

Literature review

Hariri and mehraban in analyze databases of nanotechnology through the query analysis and follow-up users navigation express the Major concern of users of information systems is information retrieval related their information needs, query used by the users is a manifestation of their information needs Results show that nanothechnology databases users using three methods to retrieve information needs: search engines, referral sites and directly use. In the used directly Bounce Rate have been lower and more pages have been visited. The average length of query is. And easier search 3.36 strategies are used to retrieve information. (Hariri and Mehraban, 2014)

Tsuyoshi et. al described a method for clarifying users' interests based on an analysis of the site-keyword graph. The method is for extracting sub graphs representing users' main interests from a site keyword graph which is generated from web log data. (Tsuyoshi et al, 2006). Jamali (2006) introduced CIBER's Virtual Scholar research program. The main aims of the program are to investigate the traits and characteristics to the virtual scholar, who has become a major efficient evidence- based study of the digital scholars' information-seeking behavior. To achieve these aims, CIBER has developed a methodology, which is termed Deep Long Analysis (DLA). DLA is more sophisticated form of transaction log analysis. He stated that researches like analysis of log files lead to develop scientific communication. Koutsoupias analyzed the log file using statistical analysation method and provided a tool for a better understanding and interpretation of the preprocessed statistical results produced from web log data, (Koutsoupias, 2002). The paper Kohavi explains about web server log files, problems of dealing with log data, lessons and metrics based on e-commerce, deficiencies of web server and presents statistics to overcome the issues. (Kohavi, 2001) Jansen and Spink provide an overview of Web-search transaction-log analyses (Bernard et al, 2006). Hofmann et al. express the need for semantic enrichment of queries with **annotations of language tags, named entities, and links to structured sources of**knowledge. (Hofmann et al, 2009) In 2004, Berendt et al ,had already argued ingeneral for further integration of the fields of Semantic Web and Web usage mining. (Berendt et al, 2004)

Research findings

As can be seen in table 1. The more pages that users have visited the site Marvdasht University is concerned to visit the home page and the second page Hot Hits is Chart educational system and the top used search engine Google and Yahoo the least by visitors.

Table 1. Summary Activity

Hits	Page Views	Visitors
The most high-traffic days	2013.13.4	75567
Most high-traffic Month	2013.6	286008
Least-traffic Month Most	2013.2	12491
First page Most visited	Hompage	641309

Secod page Most visited	educational system	169548
Third page Most visited	Educational chart	48031
Most visitors Login	http://www.google.com/url	85498
Total Web crawler Google	2291427 kB	38912
Total Web crawler Yahoo	28400KB	275

According to the Figure 1, the most visited website in the Marodasht University hour s at 22 pm and 13 pm has the fewest hits.

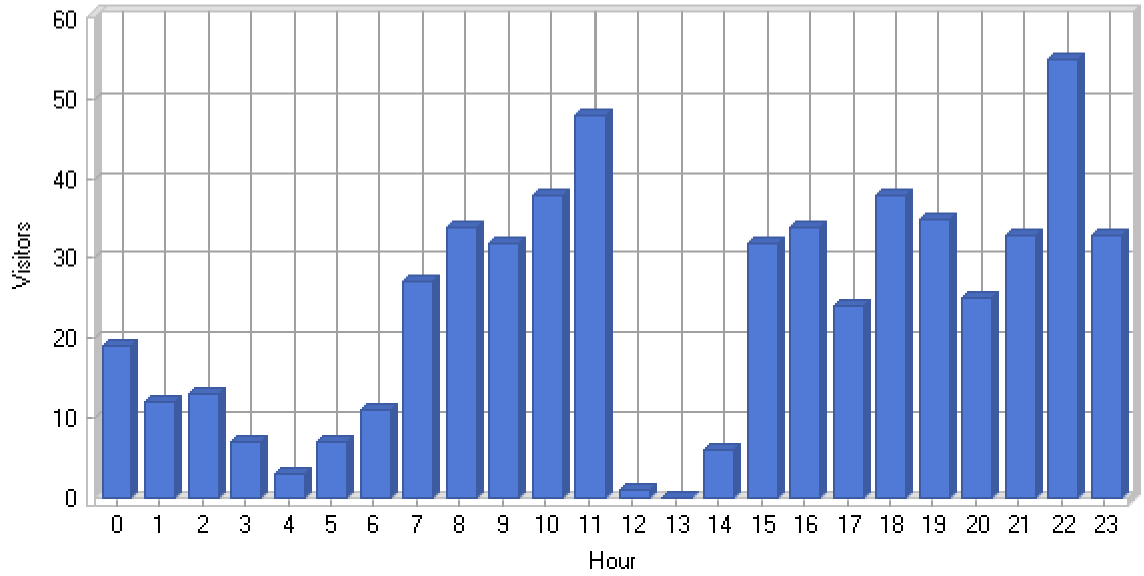


Figure 1: Activity by Hour of Day

Figure 2 shows that the highest university Site visit on Saturday, Tuesday and Wednesday have been done.

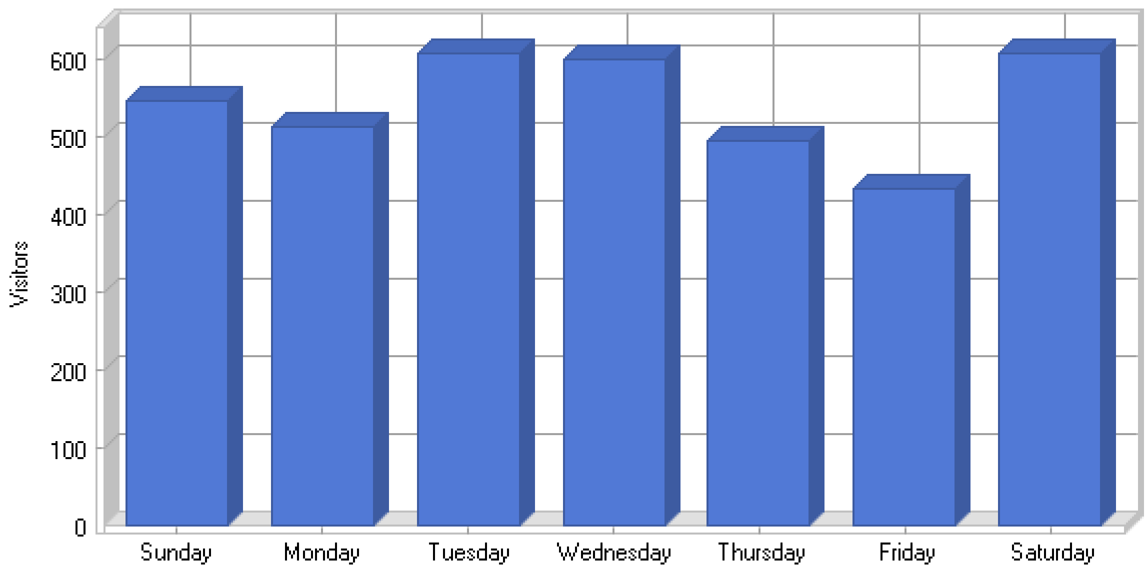


Figure 2: Activity by Day of Week

The most Requested files related to html files.

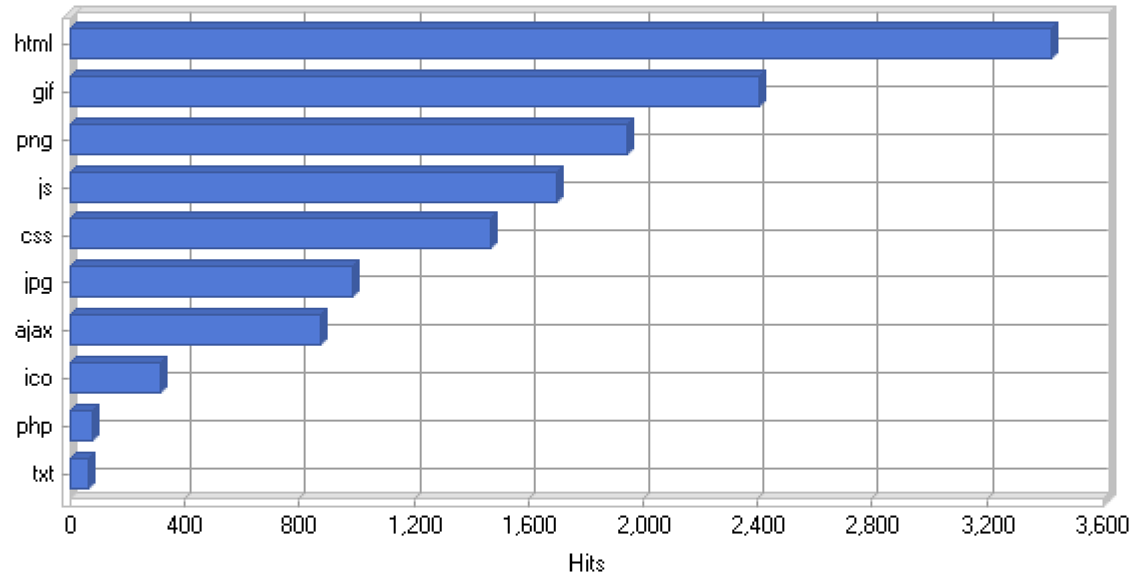


Figure 3: Most Requested File Types

Figure 4 shows that the top used search engine Google and Yahoo the least by visitors.

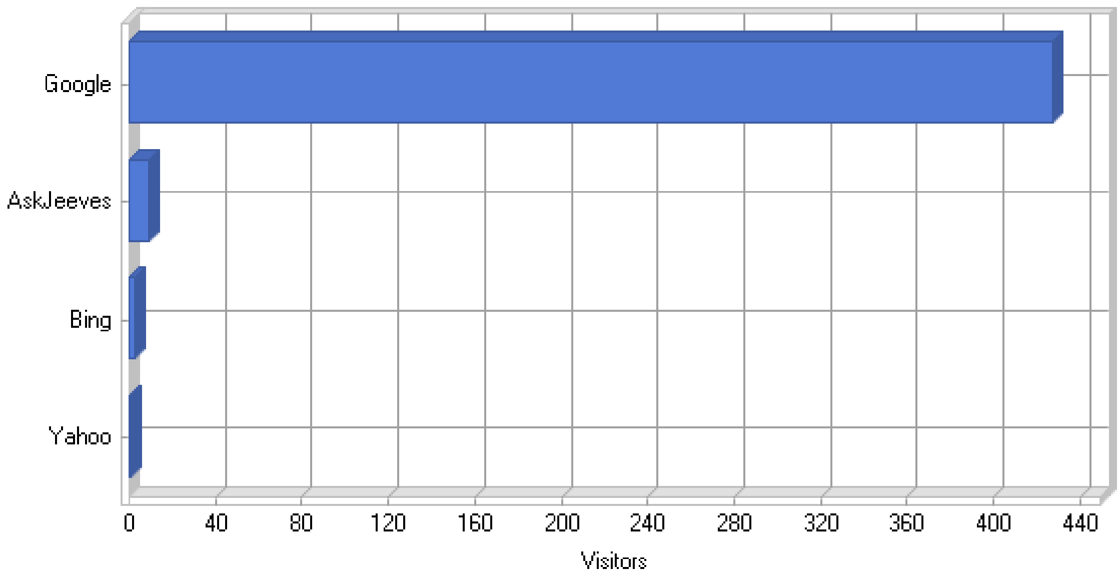


Figure 4: Top Search Engines

As can be seen in Figure 5 in the country visits site Marvdasht University, Iran has most visited.

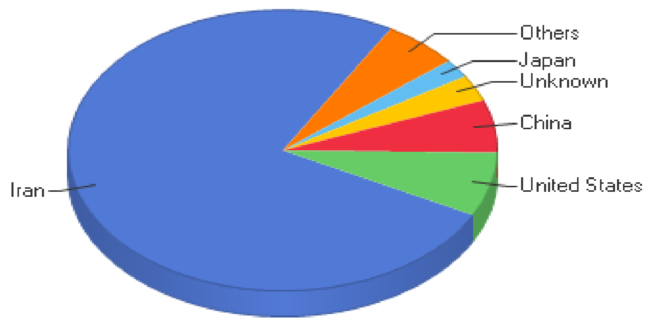


Figure 5: Most Active Countries

According to Figure 6 Firefox Browsers is most used by visitors.

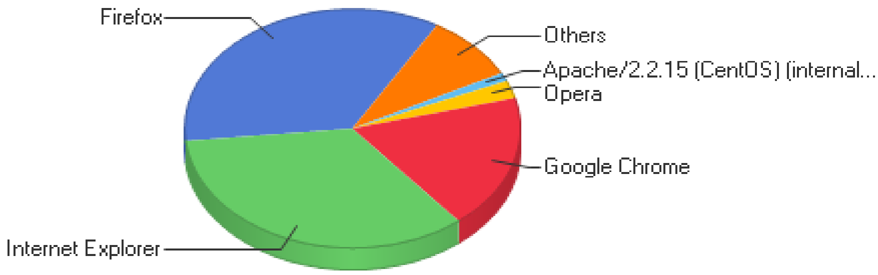


Figure 6: Most Used Browsers

As can be seen in Figure 7, the most widely used operating system is Windows 7.

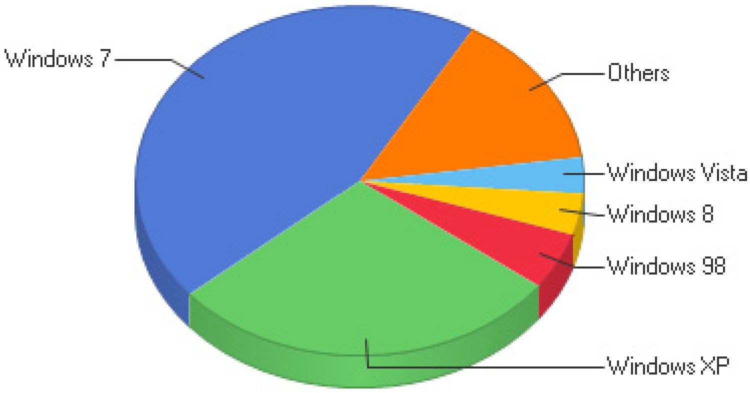


Figure 7: Most Used Operating Systems

Conclusions and Discussion

Transaction-log analysis is a valuable and often used method to study user's search behavior on the Web. It is popular for being a non-intrusive way to study large amounts of usage data. By analyzing these files can be expected to increase the performance of Web links , which have a positive impact and are very useful for web designers in order to increase user satisfaction to help.The main focus Log in user behavior and behavior helps to know what happened.

Web log transaction analysis is one of Webometrics methods. Users' feedback can be gained using this method. In this research, web log files of Islamic Azad University Marvdasht branch were analyzed. The research purpose is to achieve user behavior by using data mining technology, data were collected during six months in the form of log file in the Islamic Azad University Marvdasht branch website.. The research method is transaction log analysis.This study was performed through the analysis of the interaction between users and the website transaction files.

Research population included pages of the website visitors Islamic Azad University Marvdasht branch during 6 months (from February 2013 to July, 2013), they have visited.University website to get the required data from the server and then use the Excel for filtering unrelated data were analyzed. The results of the 5 million has been selected because it is presented as a percentage, has high level of flexibility. The final analysis is based on 500 data are expressed as the percentage. Results showed that more pages that users have visited the site Marvdasht University is concerned to visit the home page and the second page Hot Hits is Chart educational system. Most traffic has been from inside the country.

Results show that Islamic Azad University Marvdasht - branch website users using three methods to retrieve information needs: search engines, referral sites and directly use. Most visitors use of direct references to the homepage.

27 Percentage of hits has been via search engines. The top used search engine Google and Yahoo the least by visitors. 6 percent of Visitors access to the university website through the references of other Web pages. The most visited website in the Marvdasht University hours at 22 pm and 13 pm has the fewest hits.finding shows that the highest university Site visit on Saturday, Tuesday and Wednesday have been done. The most Requested files related to html files. In the country visits site Marvdasht University, Iran has most visited. Firefox Browsers is most used by visitors**.** Most widely used operating system is Windows 7.

References

Andrews J. H. {?}.A Framework for Log File Analysis."http://citeseer.nj.nec.com/159829.html.

Bernard, J. (2006). Search log analysis: What it is,what's been done, how to do it. Library and Information. Science Research, 28(3):407 - 432.

Bernard J and Spink,A. How are we searching the World Wide Web? A comparison of nine search engine transaction logs. InformationProcessing and Management, 42(1):248 - 263.

Berendt, B& et al (2011). Usage analysis and the web of data. SIGIR Forum,45(1):63-69, May 2011.

Berendt , B, Stumme, G. And Hotho, A. (2004). Usage mining for and on the semantic web. Data Mining: Next Generation Challenges and Future Directions, pages 461-480.

Hariri , N**.** Mehraban, S. (2014) Search strategies in nanothecnology databases: log analysis. http://Jipm.irandoc.ac.ir.

Hofmann ,K,& et al. (2009). A semantic perspective on query log analysis. In Working Notes for CLEF, 2009.

Hollink, L . Mika ,P . Blanco,R. (2013). Web Usage Mining with Semantic Analysis. International World Wide Web Conference Committee (IW3C2).WWW 2013, May 13–17, 2013, Rio de Janeiro, Brazil.2013 ACM 978-1-4503-2035-1/13/05.

Jaideep, S & et al.(2000).“Web usage mining: Discovery and Applications of usage patterns from web data” SIGKDD Explorations- vol-1,issue-2 Jan 2000 pages 12-33.

Jamali, H. (2006) .The virtual scholar: the results of research program. Informology. 4 (1, 2). Autumn & Winter. PP: 132-150.

Kohavi (2001). “Mining E-Commerce Data: The Good, the Bad, and the Ugly” *KDD 2001*, Aug 26-29, San Francisco, CA. Copyright 2001 AC.

Kosala, R. Blockeel, H. (2000). Web mining research: a survey,SIGKDD: SIGKDD explorations: newsletter of the special interestgroup (SIG) on knowledge discovery & data mining, ACM 2 (1),1–15.

Koutsoupias, A. (2002). “Exploring Web Access Logs with Correspondence Analysis” 2nd Hellenic Conf. on AI, SETN-2002,11-12 April 2002, Thessaloniki, Greece, Proceedings, CompanionVolume, pp. 229-23.

Lizhen, L. Chen, J. Song, H. (2002) . “The Research of Web Mining”, Proceedings of the 4th World Congress on Intelligent Control and Automation, June 10-14, Shanghai/China, 2002.

Martin Kurth. The limits and limitations oftransaction log analysis. Library Hi Tec, 11(2):98-104.

Naresh, B. (2003).“Web usage Mining and Pattern Discovery “a Survey Paper” –Dec 8 2003. http://citeseer.nj.nec.com/159829.html.

Ronald E. Rice and Christine L. Borgman. (1983) The use of computer-monitored data in nformation science and communication research. Journal of the American Society for Information Science, 34(4):247-256.

Tsuyoshi, M. and Kota, S. (2006). "Extracting Users Interests from Web Log Data" roceedings of the 2006 IEEE/WIC/ACM International Conference of Web Intelligence (WI 2006 MainConference Proceedings) (WI'06) 2006 IEEE.

Use of Six Sigma Concept in University Libraries: A Case Study of Fars province Medical Sciences Library University

Marzieh Yari Zanganeh* and Sedigheh Mohammad**

*Department of Knowledge and Information Science, Marvdasht Branch, Islamic Azad University, Marvdasht, Iran
myzanganeh@yahoo.com

**Assistant Professor, department of Knowledge and Information Science, Science and Research Branch, Islamic Azad University, Tehran, Iran

Abstract

This paper discusses the Six Sigma applications in the academic libraries especially University libraries This research aimed to investigate the perspectives of librarians in Fars province Medical Sciences Library University about the performance of Six Sigma Principles in libraries of university.This was a descriptive survey in which a questionnaire was used for data collection. The study population consisted of 130 librarians of Fars province University of Medical Sciences Bachelor's degree or higher in librarianship in 2013-2014. 107 questionnaires were returned. This research was analyzed using descriptive (number and percentage) and inferential statistics (one-sample t test for hypotheses) with SPSS software was used. For verifying reliability of the questionnaire was also estimated 83% using Cronbach's Alpha.Result In comparison to the average of six sigma in the library community show that, the average scale of one percent ($p<0.01$) there is a significant difference. This means that the use of Six Sigma in the libraries evaluated are significantly higher than average.

Keywords: Six Sigma applications, Library, Medical Sciences university

Introduction

"As technology moves more and more into the very fabric of our existence, the real-time existence of consumers and businesses and the economy, the reliability of systems will need to reach the level of dial-tone. Consequently, the techniques used to develop systems, the quality of these systems, and the demands on the performance of these systems will all need to be higher than anything we've ever imagined. This, in turn, means that company stock prices will start to be impacted by project slippage. Project slippage is going to start taking out companies**.** Six Sigma is a Greek word and in its English version implies standard deviation. Six Sigma is known by **σ** sign. Six Sigma means the system for removal of defects in the present processes and providing customers with expected specialized products and services. In simple language Six Sigma means a standard for measurement for six deviations from the mean. It is a standard to completely finish the defects in any constituent in quality. This is world reputed quality system. .(Ulhe, et al.2011)Six sigma was originated in 1980s at Motorola as a methodology of implementing TQM. The credit of inventing six sigma is to Bill Smith, an engineer at Motorola. Later Motorola established Motorola University with the programmes like black belt, yellow belt to train the people in six sigma methodology. Today these programmes are running worldwide through various organisations. Service organisations such as healthcare and finance have been implementing six sigma and are registering benefits. The breadth of applications is now expanding to other services including call centers and human resource and product support services. (Kaushik, et al. 2007)

In library and information centres the first and foremost objective is to satisfy the need of its users. To achieve this goal proper library management based on scientific principles is very important. Six Sigma is a business management strategy is to improve the quality of process outputs by identifying and removing the causes of defects and minimizing variability in manufacturing and business processes. Therefore, the Six Sigma being the tool for assessing the quality as well as the problem solving tool for corporate sectors may be applied in library management also. (Sevukan, 2011)

Library Services And Six Sigma Application

In the IT influenced environment, libraries are extremely under-pressure to justify their existence as well as their importance for the organisation. For this, libraries must satisfy user's need as well as meet their expectation. Moreover, libraries should constantly strive to provide quality services to users by cutting the costs.Six sigma is a method for improving quality by reducing errors that result in quality service with reducing costs. Using six sigma libraries can improve their service to users by reducing defects and minimising cost involved in library services. This will satisfy users as well as the funding organisation. (Agrawal, 2012) Library is a place where knowledge is discovered. Driven by this philosophy, the present study focused on using the DMAIC (define, measure, analyse, improve, control) methodology to improve the efficiency of a local library. Thefollowing process improvement steps were taken for this pilot study (Kaushik, et al. 2007)

Research hypotheses

1 - the real focus of the research community on the client libraries is desirable.
2 - management based on facts and data libraries of the study population is desirable.
3 - Focus on processes in the libraries of the study population is desirable.
4 - action management of the libraries of the study population is desirable.
5 - Participation and Collaboration limitless libraries of the study population is desirable.
6 - Moving Towards Excellence and tolerate failure in the libraries of the study population is desirable.

Literature review

Dong-Sug Kim (2010) on advantages and disadvantages of Six Sigma implementation in Sungkyunkwan University. Sungkyunkwan University has applied 'six sigma' in every department of university and library of this university was one of them. Researcher collected data through interview and questionnaire and analysed it using qualitative as well as quantitative techniques. Researcher observed some positive opinion of six sigma implementation, i.e. making work method scientific, increasing process capacity, turning subjective knowledge into a formal format,etc. while the negative opinions include lack of time of participation, lack of interest in employees, poor standardization, and difficulty with defining work process.

Yong kim, et al (2009) applied six sigma in library acquisition process and the results proved that services of the library acquisition was good and better after implementing six sigma tool.

Susan Kumi and John Morrow (2006) observed that six sigma suggested the defects in self-service at Newcastle University Library. The library benefited from six sigma not only in that it achieved its goal of increasing self- service percentage, but it also provided them with a strong method of addressing a problem accurately and speedily in a systematic way.

Antony (2004) defines it as “a strategy that seeks to improve the quality of processes through identifying and removing the causes of defects by focusing on outputs that are critical to customers”. Thus, in one way six sigma refers to a measure of process consistency and aims at achieving the same.

Benefits of Six Sigma System in libraries

1. Since increased type of services in library is available there is satisfaction to the readers;
2. This system exerts influence on those who serve and they give better type of service;

3. Since excellent service is available the figure of reader increase, yielding financial benefits;
4. Due to excellent service and quality the readers are satisfied. (Ulhe et al., 2011)

Methodologies of Six Sigma Tool

As a disciplined process, six sigma provides two standard process models, i.e. DMAIC (Define, Measure, Analyse,Improve, and Control) and DMADV (Define, Measure,Analyse, Design, and Verify). DMADV is aimed at development of a new product or process, while DMAIC is for improvement of existing process or product. Thus, appropriate strategy to be chosen by the six sigma team. (Agrawal,2012)

DMAIC which is best suited to the library environment. The method insists to have continuous assessment, improvement, and guide to bring out excellent services to library users. There are five stages in DMAIC methodology to improve the quality, service, and resources of the library .First emphasize is laid on "Define" the problems, the opportunity, the process, the projects, the goals and the users. The second one is "Measure", which helps you to decide current level, current process and decide customer needs and requirements. "Analyze" is the step which guides you to decide the origin and source of the defects. The fourth step "Improve" is to improve the process by eliminating defects / performance / current procedure / standard of work. At last the finest step is "Control", which makes you to look and take control all the above acts. DMAIC cycle method should be repeated again and again for continuous improvement. (Sevukan, 2011)

Research method

This study is a descriptive survey in which a questionnaire was used for data collection. The study population consisted of 130 librarians of Fars province University of Medical Sciences Bachelor's degree or higher in librarianship in 2013-2014. 107 questionnaires were returned. This research was analyzed using descriptive (number and percentage) and inferential statistics (one-sample t test for hypotheses) with SPSS software was used. For verifying the face validity of questionnaire, it was given to other professors of Management and Information Sciences and knowledje and their suggestions were applied and the reliability of the questionnaire was also estimated 83% using Cronbach's Alpha.

Research findings

In this study, data collected from the questionnaire in two parts: demographic information, participants will analyze and test the hypotheses. According to Peter pande and Lawrence Halp six sigma principles (Pande & Holpp, 2001): real focus on on the client, focus on process, management based on facts and information, Action management , Participation and Collaboration limitless and moving Towards Excellence and tolerate failure in the libraries of the study population in 6 hypothesis were examined.

Table 1: Demographic information of the study population

demographic information		Frequency	Percent
sex	femail mail	87 20	81.0 19.0
dicipline	librarian nonlibrarian	98 9	91.8 8.2
Organizational post	staff manager	93 14	86.7 13.33

In this study, 107 librarians participated that 81% femail and 19% of them were males. 91.8% of them have a degree librarianship and 8.2% were non-Librarianship. 86.7% of them employees and 13.3% were libraries Managers (Table 1)

Table 2: Comparison of the six sigma principles in the libraries of the study population with desirable level

six sigma principles	Test Value =						
	t	df	Mean	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
						Lower	Upper
Real focus on the client	15.726	29	3.8000	.000	3.80000	3.3058	4.2942
Management based on facts and information	16.551	29	3.0667	.000	3.06667	2.6877	3.4456
Focus on process	24.144	29	3.9000	.000	3.90000	3.5696	4.2304
Action management	18.784	29	3.5667	.000	3.56667	3.1783	3.9550
Participation and Collaboration limitless	17.299	29	3.8667	.000	3.86667	3.4095	4.3238
Moving Towards Excellence and tolerate failure	21.153	29	3.6000	.000	3.60000	3.2519	3.9481

Comparison of Six Sigma principles in the libraries of the research community with desirable level, six hypotheses were proposed in this study. As seen from the data in Table 2 , test results in six cases showed that there were significant differences in the mean scale,This means that the use of Six Sigma in the libraries evaluated are significantly higher than average.

Conclusion

The IT Infrastructure Library and Six Sigma should be viewed as complementary systems that leverage one another in the common goal of world-class IT performance. Organizations should use this powerful combination to bolster their competitiveness.Six sigma was introduced for manufacturing process, but for more than two decades its implementation is also seen in service industries. Though, it is not applied in Iranian libraries on a wide scale, it would not be justifiable to say that this process is not applicable in libraries. Library is the organisation which needs to focus on quality of service and user satisfaction. Six Sigma is generally used in manufacturing sectors to minimize the wastages and to assure this quality in such a way the same can be implemented in libraries to develop the process and improve the standard of the library to satisfy the users. It insists on continuous improvement and development of the library as well as library staff members. The ultimate goal of the library is to satisfy its users. This can be achieved only by applying and experimenting new tools and techniques available today for libraries. The past case studies show that it has a good applicability in libraries also and can produce good results.

References

Agrawal, P. R.(2012) Application of 'Six Sigma' in Libraries for Enhancing Service Quality International Journal of Information Dissemination and Technology | October-December 2011 | Vol.1 | Issue 4.

Antony, J. (2004). Some pros and cons of six sigma: an academic perspective. *The TQM Magazine, 16*(4), 303 -306.

Kim, D.-S. (2010). Eliciting success factors of applying Six Sigma in an academic library. *Performance Measurement and Metrics, 11*(1), 25-38.

Kumi, S., & Morrow, J. (2006). Improving self service the six sigma way at Newcastle University Library. *Program:electronic library and information systems, 40*(2), 123-136.

Kaushik, C. (2007). Six Sigma applicaiton for library services. *DESIDOC Bulletin of Information Technology, 27*(5), 35-39.

Pande, Peter; Holpp, Lawrence. *What Is Six Sigma?* (2001) McGraw-Hill Osborne Media.

Sevukan. R. Suresh. N. Madasamy. R. (2011). *Application of six sigma tool in library management: a bird's eye view*. Department of Library and Information SciencePondicherry University, Puducherry – 14www.smartlibrarians.in.

Ulhe, U. S., & Patil, S. K. (2011). Application of SixSigma Concept in University Libraries, a Key to Success:A case study of Jayakar Library, University of Pune (India). *Asia Pacific Journal of Library and Information Science, 1*(2), 134-142.

Yong Kim et al., (2010). A Six Sigma-based method to renovate information services Focusing on information acquisition process. *Library Hi Tech*, Vol.28, No.4, pp 632-647.

How is scientific research reported in newspapers? – Comparison between press releases and two different national newspapers in Japan

Masaki Nishizawa and Yuan Sun

2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo, Japan
nisizawa@nii.ac.jp, yuan@nii.ac.jp

Abstract

We analyzed the frequency at which press releases from universities form the basis of scientific articles in newspapers. The number of universities that have been publishing press releases has been growing. Moreover, the number of newspaper articles on the university sector has also been growing. In this study, based on data of press releases and two major Japanese national newspapers, we investigated what the relations between press releases and newspapers are and how differently scientific research is reported in the two newspapers.

Introduction

How are the results of research institutions reported to the wider public? Press releases from organizations, such as universities, have been increasing in number partly as a means to show the accountability necessary to secure research funds and partly as a way to increase enrolment. Moreover, research on press releases related to top-tier Japanese universities has shown a generally increasing trend, and the number of organizations with active press release programs has grown rapidly in recent years (Nishizawa M. and Sun Y. 2012). It has also been shown that academic articles related to universities in newspapers are increasing with the increase in press releases (Nishizawa M. and Sun Y. 2013).

Press releases are public relations media, but they only reach the wider public after being published in a newspaper or other media under the control of another organization. We used data from two major Japanese national newspapers from 2007 to 2012 to see if there were any differences between the articles reported in them, and if so, how they differed. We investigated the circumstances under which press releases would be used in newspaper articles.

Data

Name identification database

We previously developed a name identification database of author affiliation organizations with about 20,000 records. For our research, this database was augmented with about 55,000 records with company name identifications. In particular, a database with data on 40,000 Japanese companies (List of Companies) belonging to Toyo Keizai, Inc. (Toyokeizai Online, 2014) and a DCS-Organization name dictionary published by Nichigai Associates, Inc. (Nichigai Associates, 2014), were used to find the exact and parent-child company names. A total of 76,739 records from the name identification database were used in the analysis presented below.

Press releases

Press releases are announcements from organizations that are released to the public via newspapers or other media. University-related press releases can function as announcements of (1) research results, (2) industry-university cooperation including tie-ins with specific companies, and (3) reports on products received by the university from companies. We used the Nikkei Telecom 21 search service (Nikkei Telecom 21, 2014) and found 10,120 press release titles that contained the word "university" between 2005 and 2012. Our search involved using the full text and title, so releases in which the word "university" is not included in the title were also gathered. The total number of press releases during this period was 235,209, so 4.3% included the word "university".

We extracted organization names included in the title by using the name identification database and a Japanese morphological analysis system (Chasen); we extracted 12,855 organizations from 10,120 titles.

News Paper

We have been using corpora of three major Japanese national papers (Nichigai Associates, Inc., 2014). For this paper, we present our results for the Yomiuri Shimbun and Mainichi Newspapers (articles published between 2007 and 2012, respectively). The organization names were extracted from the article's text body using the name identification database described previously. Table 1 lists the number of articles from the two major Japanese newspaper and press releases.

Table 1: Number of articles from two major Japanese newspaper and press releases

News paper	*2007*	*2008*	*2009*	*2010*	*2011*	*2012*	*2013*
Mainichi	297568	296991	288088	283373	285608	276508	257672
Yomiuri	343256	333208	316602	310530	312937	309958	299824
Press Release	28611	29947	31110	29585	29830	27031	25099

Correlation of press releases and newspaper articles

Methods

To investigate the correspondence of university-related press releases and the content of newspaper articles, we analysed the correspondences of organization names appearing in the full text of newspaper articles or their titles.

About 51% of the articles identified an organization by name. These articles were then compared with the press releases. To shorten computation time, we limited the date of publication of the newspaper articles to within 15 days before and 45 days after the press releases. After extracting keywords from each title and text body of the article, using Chasen, the cosine distance (*Sim(A,B)*) to each keyword of the press releases was calculated for a pair of articles in which the same organization name appeared.

$$Sim(A,B) = \frac{\sum_{i=1}^{n} W(A_i) \times W(B_i)}{\sqrt{\sum_{k=1}^{n}(W(A_k))^2} \times \sqrt{\sum_{j=1}^{n}(W(B_j))^2}}$$

Here, A and B are the keywords of newspapers and press releases, respectively and W is the weight of each keyword, which uses the natural logarithm of the frequency N of a keyword. The correspondences of the content on the basis of the cosine distance and the difference between the printing date of the article and the date of the press release were manually checked. The following *corres_index* was introduced as a standard for measuring the degree of correspondence of press releases and newspaper articles from the cosine distance and time difference (*date_diff*).

$$corres\_index = Sim(\text{headline, PR})/0.4 + Sim(\text{textbody, PR})/0.3 + 0.5/diff\_index;$$

$$diff\_index = \text{int}(date\_diff/2) + 1 \quad :(date\_diff >= 0)$$
$$= \text{int}(14/date\_diff) - 1 \quad :(date\_diff < 0)$$

Figure 1 shows the frequency distribution of the cosine distance for all investigated (Fig.1-1) and correspondent (Fig.1-2) instances for press release title vs. newspaper headline: *Sim(headline, PR)* and newspaper text body: *Sim(textbody, PR)*. To arrange influence of the cosine distance on a headline and the text body, 0.4 of *Sim(headline, PR)* was standardized to 1 and 0.3 of *Sim(textbody, PR)* was standardized to 1, respectively. These constants for standardization originate from the averages of the cosine distances of the correspondent reports having been 0.39(PR-Headline) and 0.33(PR-Textbody).

To select newspaper articles in agreement with the content of the press releases, manual checking was done using this corres_index index. The frequency distributions of the *corres_index* for the correspondent ratio for the Yomiuri Shimbun and Mainichi Newspapers are shown in Fig. 2. If *corres_index* has a value of 2.5 or more, the correspondence rate is 90% or more. To the newspaper article of press releases and two or more correspondence candidates, this *corres_index* is in the top position in almost all examples, and the checking efficiency improved greatly. However, optimization of this index is for future work.

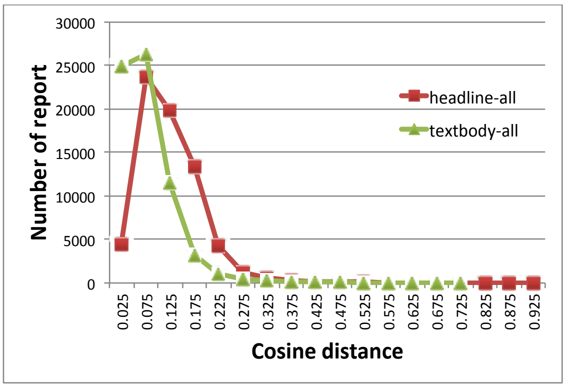


Fig. 1-1: All investigated instances

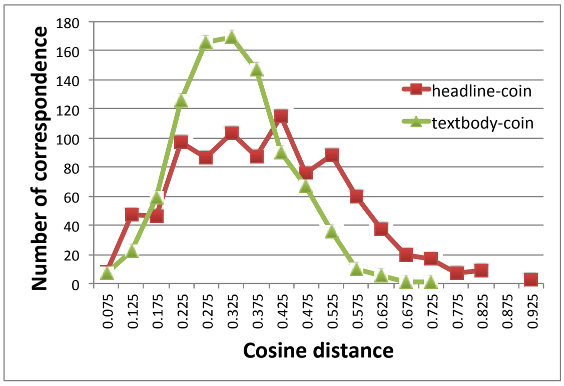


Fig. 1-2: Correspondent instances

Figure 1: Frequency distribution of the cosine distances for all and correspondent instances (Yomiuri Shimbun)

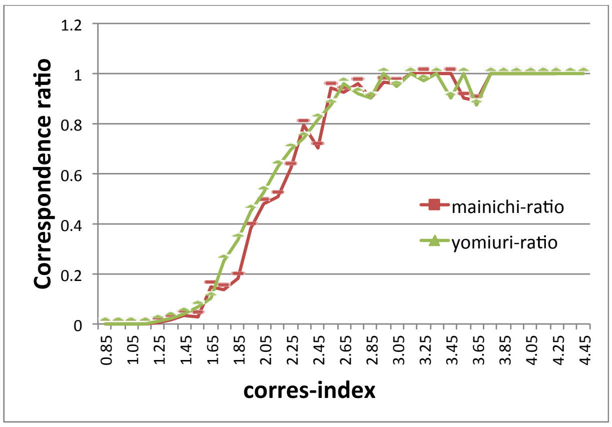


Figure 2 Frequency distribution of *corres_index* for correspondence ratio (Yomiuri, Mainichi)

Result and Discussion

Press releases

The 2012 data were added to our previous results (Nishizawa M. and Sun Y. 2013) (Table 2). Although detailed results (Nishizawa M. and Sun Y. 2013), such as classification by the last verb in a press release title, have not yet been obtained, the growth of Kyoto University is conspicuous from the 2012 data..

Table 2: Ranking of organization name instances appearing in titles (Univ. and InstN sectors)

Rank	Organization	sector	total	2005	2006	2007	2008	2009	2010	2011	2012
1	Univ. of Tokyo	Univ	571	22	20	21	16	38	86	186	182
2	RIKEN	InstN	472	6	17	68	90	73	68	79	71
3	Tohoku Univ.	Univ	414	7	21	33	45	58	63	83	104
4	AIST	InstN	221	31	28	34	40	18	27	18	25
5	Kyoto Univ.	Univ	152	5	3	7	14	5	10	37	71
6	Keio Univ.	Univ	133	9	5	19	11	20	12	14	43
7	Hokkaido Univ.	Univ	107	5	3	2	7	6	31	31	22
8	NICT	InstN	105	16	14	9	13	13	18	12	10
9	Osaka Univ.	Univ	65	9	3	3	6	8	7	10	19
10	Kyushu Univ.	Univ	57	5	4	6	6	2	6	10	18
11	Waseda Uinv.	Univ	52	5	6	7	13	7	5	5	4
12	Nagoya Univ.	Univ	44	5	4	3	4	4	7	8	9
13	NIPS	UnivI	42				1	12	12	9	8
14	Tokyo Inst. Tech.	Univ	34	5	3	2	4	2	5	4	9
by sector		Univ	2516	161	162	204	212	259	351	536	631
		InstN	916	64	69	122	160	116	132	124	129
		Corp	8462	1103	1063	1123	1060	1143	1012	1037	921
		Other	374	25	28	34	31	30	44	81	101
		unknown	565	123	83	77	54	53	63	50	62
		Org. Total	12833	1476	1405	1560	1517	1601	1602	1828	1844
	search by "Univ"		10120	1154	1097	1216	1209	1244	1299	1429	1472
	all press release		235209	29311	29784	28611	29947	31110	29585	29830	27031

Correlation of press releases and two national newspaper

We investigated the correlation of 10,120 press releases and reports of two Japanese national newspapers. Table 3 lists the ranking of organization names appearing in the correspondence report published at Yomiuri Shimbun. We found 635 correspondences with press releases. The number of different organizations that appeared in this correspondence report was 793.

Table 3: Organization ranking of correspondence reports of Yomiuri Shimbun and press releases.

Rank	*Organization*	*sector*	*total*	*2007*	*2008*	*2009*	*2010*	*2011*	*2012*
1	Tohoku Univ.	Univ	91	9	19	8	20	11	24
2	RIKEN	InstN	59	12	12	7	8	11	9
3	Univ. of Tokyo	Univ	57		1	2	12	18	24
4	Kyoto Univ.	Univ	46	1	6	1	1	14	23
5	Keio Univ.	Univ	18	2	2	1	1	2	10
6	AIST	InstN	15	6	3	3	3		
7	Hokkaido Univ.	Univ	15		2	3	6	3	1
8	Osaka Univ.	Univ	13	2	1			4	6
	by Sector	Univ	332	28	44	25	56	69	110
		InstN	84	18	15	12	13	12	14
		Corp	361	67	66	39	52	85	52
		Other	16	1	4	0	2	3	6
		All Sector	793	114	129	76	123	169	182

Table 4 lists the ranking of organization names appearing in the correspondence report published at Mainichi Newspapers. We found 465 correspondences with press releases. The number of different organizations that appeared in this correspondence report was 555. There were about 73% fewer correspondences for Mainichi Newspapers than for Yomiuri Shimbun.

Table 4: Organization ranking of correspondence reports of Mainichi Newspapers and press release.

Rank	*Organization*	*sector*	*total*	*2007*	*2008*	*2009*	*2010*	*2011*	*2012*
1	RIKEN	InstN	51	15	14	6	8	5	3
2	Univ. of Tokyo	Univ	50	1	1	2	11	19	16
3	Tohoku Univ.	Univ	40	2	9	6	7	8	8
4	Kyoto Univ.	Univ	30		4	1	1	9	15
5	AIST	InstN	23	9	4	1	5	3	1
6	Keio Univ.	Univ	17	1	1	1	2	2	10
7	Hokkaido Univ.	Univ	11		2		4	3	2
	by Sector	Univ	213	12	29	14	35	52	71
		InstN	84	24	20	9	14	11	6
		Corp	244	42	47	28	27	48	52
		Other	14	0	2	2	2	4	4
		All Sector	555	78	98	53	78	115	133

Table 5 lists the ranking of organization names appearing in the correspondence reports published at both Yomiuri Shimbun and Mainichi Newspapers. The congruous number of press releases was 240. That is, 53% to Mainichi Newspapers and is 38% to Yomiuri Shimbun. More than half of the correspondence articles in Mainichi Newspapers were also published at Yomiuri Shimbun, and it seems that the selection policy is mostly in agreement. A detailed analysis of certain areas of research, such as, the journals in which source papers were published, is for future work.

Table 5: Organization ranking of correspondence reports of both Mainichi Newspapers and Yomiuri Shimbun by year (Univ, InstN sector)

Rank	*Organization*	*sector*	*total*	*2007*	*2008*	*2009*	*2010*	*2011*	*2012*
1	RIKEN	InstN	35	9	8	5	5	5	3
2	Univ. of Tokyo	Univ	27			1	7	11	8
3	Kyoto Univ.	Univ	26		4	1		8	13
4	Tohoku Univ.	Univ	25	1	4	4	6	6	4
5	Keio Univ.	Univ	10	1	1		1	1	6
6	AIST	InstN	8	3	2	1	2		
7	Hokkaido Univ.	Univ	5		1		2	2	
	by Sector	Univ	125	6	15	8	22	34	40
		InstN	45	12	10	7	7	5	4
		Corp	118	19	28	8	11	26	26
		Other	6	0	2	0	0	2	2
		All Sector	294	37	55	23	40	67	72

Figure 3 shows the growth rate in press releases and newspaper reports from 2007 (Yomiuri Shimbun) that are correspond with press releases by year for each sector. Only the university sector is growing.

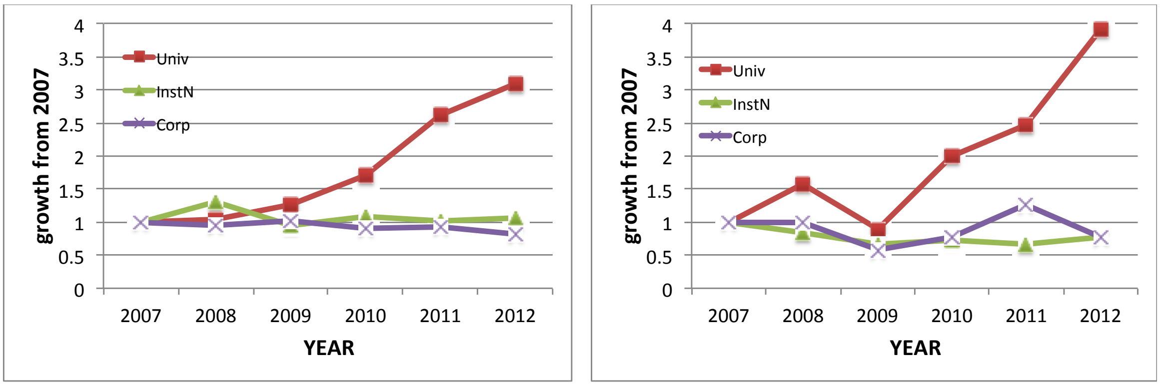


Fig. 3.1: Growth in press releases

Fig. 3.2: Growth in coincident news

Figure 3: Growth rate in press releases and correspondent newspaper reports from 2007 (Yomiuri Shimbun)

In the university sector, the number of correspondence articles in newspapers has increased about 4 fold since 2007 depending on about 3 times growth of press releases. The number of newspaper articles is also increasing along with the issuing of the press releases in the university sector, and the positive announcement to press releases has an effect on their incorporation into newspaper articles.

Summary

We investigated how the results of academic research from universities are reported to the wider public by focusing on press releases and two Japanese national newspapers. We found 635 correspondences between the title of press releases and reports from Yomiuri Shimbun and 465 correspondences for Mainichi Newspapers. Furthermore, the number of press releases by both national papers was 240.

Both the number of newspaper articles correspond with press releases and press releases increased in the university sector. This result implies that the incorporation into newspaper articles increased as a result of increasing press releases. Further detailed analysis is needed since a difference was observed in the relation between the number of announcements of press releases and that of correspondent newspaper articles from checking individual universities. As a result of introducing the corres_index index for discovering the correspondence of press releases and newspaper articles, the checking efficiency of correspondence reports improved. However, each coefficient should be optimized.

Since original research paper information is recorded in the body text of a press release, extraction of this information is also for future work. Comparison among research funds, number of papers from scientific journals, etc. is necessary to investigate other newspaper articles.

References

Nichigai associates, Inc.: *Language resources*, http://www.nichigai.co.jp/sales/corpus.html (referenced March, 2014, in Japanese).

Nikkei Telecom 21: http://t21.nikkei.co.jp/ (referenced March, 2014, in Japanese).

Nishizawa M. and Sun Y., “How are universities collaborating with industry in Japan? --- An investigation based on newspaper press releases”, in *Proceedings on COLLNET and WIS 2012*, Soul, Koria.

Nishizawa M. and Sun Y., “How well do newspapers describe scientific research? --- An Investigation into the Correlation of Press Releases and Newspaper Articles”, in *Proceedings on COLLNET and WIS 2013*, Tartu, Estonia.

Toyokeizai Online: http://dbs.toyokeizai.net/en/ (referenced March, 2014).

Research Output of University College of Medical Science, University of Delhi: A Bibliometric Study

Meera* and Surendra Kumar Sahu**

*Assistant Professor, Department of Library & Information Science, University of Delhi,
New Delhi – 110007, India
meeradlis@gmail.com

**Research Scholar, Department of Library & Information Science, University of Delhi,
New Delhi – 110007, India, Mob.: +919213911374
say2surendra@gmail.com

Abstract

This study aims to depict the research performance of University College of Medical Science (UCMS) in different areas or subfields of medical and health sciences. This study is based on a bibliometric analysis of scientific research output. The data of UCMS' research output are collected from the SCOPUS database by using different searching techniques. Some bibliometric indicators such as authorship pattern, degree of collaboration, author productivity, rank distribution etc. have been used to illustrate the research performance of researchers. A total of 2557 research papers have been published by researchers of UCMS since 1975 to November 2013.

The result of this study shows that the highest contribution of 25.6 percent of total publications is made by three authorship collaboration. The degree of collaboration is 0.92, which means most of the research works are collaborative works. USA is the most preferred country by the researchers for research collaboration. This study will assist in understanding the research pattern of UCMS in the field of medical science.

Keywords: Bibliometric, medical and health science, research output and University College of Medical Science (UCMS), Authorship Pattern, Collaboration.

Introduction

The research output is one of the indicators which determine the development of a country in terms of Per Capita Income, GDP (Gross Domestic Product), GNP (Gross National Product), and living standard of people. Research and development are in direct proportion with each other. Research output depends upon R&D expenditure, government policies, private agencies, universities and institutions. In India, Universities and institutions are playing a significant role in R&D but still they are not recognised at the international level. Indian's R&D expenditure is less than 2.5% of global investments ($1.2 trillion) and is currently under 1% of the GDP. Presently, the Indian Science Sector is undergoing significant changes. The gross budgetary support for the science and technology sector has significantly increased during the last decade. India is ranked ninth globally in the number of scientific publications. The Composite Annual Growth Rate (CAGR) of Indian publications is around 12±1% and India's global share has increased from 1.8% in 2001 to 3.5% in 2011 (Science, technology and innovation policy 2013, Government of India). In India during the last decade, the active areas of research activities were chemistry, physics, materials sciences, engineering and clinical medicine (Thomson Reuters, 2012).

To depict these different trends or current state of scientific research output in particular geographical areas or subject disciplines bibliometric analysis is used. It involves the application of statistical methods to obtain particularly scientific productivity-related information, which is necessary to the assessment, planning, and management of growth in a given scientific journal, subject or geographical area. These methods are mainly quantitative, and are also used to illustrate qualitative pictures of scientific activities. These bibliometric analysis tools are used in this study to assess the contribution of University College of Medical Science (UCMS) in the field of medical and health sciences. Medicine and health science is one of the most research productive subject areas in India. Medical educational infrastructure in the country has shown rapid growth during the last 20 years. The country has 356 medical colleges, 297 Colleges for BDS courses and 140 colleges conducting MDS courses. (National Health Profile (NHP) of India, 2012)

University College of Medical Science

University College of Medical Science (UCMS) was established in 1971 as a Constituent College of the University of Delhi. It is one of the prime medical colleges in India. UCMS is known for its quality research with its national ranking 7th based on the number of publications in Pub Med indexed journals. There are 21 departments with various medical and paramedical courses. There are about 205 faculty members and 550 Scholars and students are working in the College. Guru Teg Bahadur Hospital is the associated teaching hospital with more than 1000 beds (University College of Medical Science, 2013).

Literature Review

The pattern of research output of an institution or country can be determined by using different types of bibliometric analysis such as their publication growth, research impact, author productivity, citation counts and collaboration pattern. Maharana & Sethi (2013) conducted a study on scientific research output of Sambalpur University, India to measure the contribution and research output of university during 2007-11. They found that Chemistry is the most favoured research areas followed by Physics, Astronomy, Plant science etc. The degree of collaboration is 0.99 (nearly equals to 1) that means there is few/small contributions by single authors and Indian Institute of Technology were the most prolific institution next to Sambalpur University. Savanur & Konnur (2012) exposed quantitative growth and development of the Bangalore University (BU) in Science and Technology in terms of publication output from 1970 to 2010. The findings present that during 1996 to 2000 the growth rate of the publications was highest of all the years of the publications that is, (131.86%) publications. Subsequently there was a gradual decrease in the growth rate in the five-year blocks of 2001 to 2005 and 2006 to 2010. BU has collaborated with 27 countries and USA is the top collaborating country with 74(31.09%) of papers followed by France with 20(8.4%). Authorship and collaboration trend were towards multi-authored paper. The prolific authors were: S. M. Mayanna with 113 papers and N. Rudraiah with 101 papers. P. V. Kamath with 98 papers with the highest h and p values 21 and 26.52, respectively. Vasishta (2011) examined the contribution and impact of research output of PEC University of Technology. The research output of the PEC was increased with average annual growth rate of 131.85 per cent during 1996 to 2009. Growth in the academic research output is seen after the PEC has acquired the deemed university status. But the international collaborative research activity in the university is still very small, accounting for just 6.21 per cent share. KAO & PAO (2009) revealed a nation-wide evaluation of research performance in management for 168 universities in Taiwan. In addition to the popular indicators of SCI/SSCI journal publications and citations, the number of projects funded by the National Science Council of Taiwan was used to account for the special characteristic of the field of management. The results show that public universities, in general, performed better than private ones due to more financial support from the government. Universities with specific missions had comparable performance to general comprehensive ones. **Prathap & Gupta (2009)** investigated the ranking of Indian universities for their research output. They proposed a more rational procedure for ranking the research performance of universities by identifying indicator that combines quality with quantity. **Sevukan & Sharma (2008)** analyzed the research performance of biotechnology faculties in central universities of India from 1997 to 2006. They found that the growth of literature in biotechnology has steadily increased; BHU is leading from the front with 42.55%; two-authored publications predominate amongst the pattern of authorship and the application of Bradford's law does not fit into the literature analyzed. Dhawan & Gupta (2007) examined the characteristics of India's physics publications output. The study found that India's physics related contribution is significantly high (86 per cent) in Science Citation Index (SCI) - covered journals, of which 26.4 per cent was in high-impact journals (IF = 1.5). The physics research activity is led by a select number of institutions in the country. The academic sector, being the biggest of all the sectors in terms of participating institutions, made the largest contributions to the physics output, followed by R&D sector, industrial sector, and government sector. However, the share of academic sector in high-impact journals was at second rank; the R&D sector topping the list. R&D sector also exceeds all other sectors in terms of publication output per institution.

Objectives

The main objectives of this study are

1) This study aims to depict the research performance of UCMS in different areas or subfields of medical and health sciences by measuring the growth rate of research output.
2) To study the subject dispersion of medical science in order to ascertain its interdisciplinary and multidisciplinary character.
3) To identify the authorship pattern and the degree of collaboration with the help of different measures.
4) To identify the most prolific authors, highly cited authors and distribution of output by citations.
5) To determine the most preferred journals by the researchers and faculty members of UCMS.

Research Methodology

This study is conducted to analyse the research output of University College of Medical Science (UCMS), University of Delhi, New Delhi, and its collaboration with other. It is based on a quantitative analysis of scientific research output published as journal articles, letter, review, conference paper, short survey, book chapter, etc. in the discipline of medical and health sciences. The data for the study has been drawn from SCOPUS database. SCOPUS is an international multi-disciplinary database indexing over 21,000 titles from more than 5,000 publishers, including 20,000 peers reviewed journals, 390 trade publications, 370 book series, and 5.5 million international conference/ seminar papers. Scopus has a worldwide coverage, of which more than half of the Scopus contents originate from Europe, Latin America and the Asia & the Pacific Region. (Scopus, 2013)

The research output data of UCMS is collected by using different searching facilities provided by the SCOPUS database. … A total of 2557 research papers has been collected from the beginning (1975) to November 2013. In the study, advanced bibliometric indicators are used to assess the research output and author productivity. For the analysis of data following indicators has been used:

Collaborative Index is the number of authors per paper

$$CI = {}^{Np}/_{Na}$$

Where, Np = Number of papers

Na = Number of authors

Degree of collaboration

$$DC = \frac{Nm}{Nm + Ns}$$

Where, *DC* = Degree of collaboration

Nm = Number of multi-authored publication published during the year

Ns = Number of single-authored publication published during the year

Collaborative Coefficient

$$CC = 1 - \frac{[F_1 + {}^{F_2}/_{2} + {}^{F_3}/_{3} + \cdots {}^{F_k}/_{k}]}{N}$$

Where, F_1 = single authored papers

F_2 = double authored papers

F_3 = three authored papers

N = Total number of publications

Data Analysis and Results

Document type wise distribution of research output

Table 1 shows the distribution of research output by document types. The research outputs of researchers of UCMS are published in 11 different document types. The researchers are mostly preferred to publish their research work in journals. Out of a total of 2557 publication, 1914 items are published in research journals as articles which comprised 74.85 percent of the total. The next document type is Letter, which was the second preference of researchers, with a number of 296 items (11.58 percent). The review was the next preferred category of document type which covered 164 items (6.41 percent). The other remaining categories are covered only 7.16 percent of total publications.

Table 1: Document type wise distribution of research output

SN	Document Type	No. of Publication	Percentage
1	Article	1914	74.85
2	Letter	296	11.58
3	Review	164	6.41
4	Conference Paper	35	1.37
5	Article in Press	33	1.29
6	Editorial	33	1.29
7	Note	30	1.17
8	Undefined	26	1.02
9	Short Survey	22	0.86
10	Erratum	3	0.12
11	Book Chapter	1	0.04
	Total	2557	100

Note: Number of results: 2557

Year wise distribution of research output

Fig. 1 depicts the distribution of research output in the duration of every three years from 1975 to 2013. The figure clearly presents continuous increase, which can be categorise in three types of growth, in research publications, i.e. low growth, medium growth and high growth. The period from 1975 to 1989 is known as low growth period, in this period of 15 years the number of publication was increased from 8 to 48 only. The medium growth period is started from 1990 to 2007 in which the number of publications are reached at 313 from 48. In last six years 300 papers are added in the research output which indicates fast growth and highly productive period.

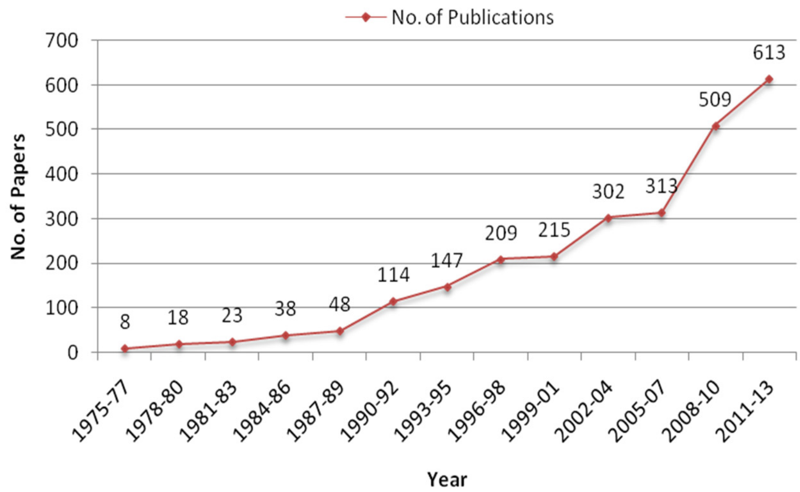


Figure 1: Year wise distribution of research output

Note: Number of results: 2557

Subject-wise rank distribution of publication

Subject-wise distribution of publications was analyzed and presented in figure 2. The papers are classified in 22 subject areas, which are directly collected from SCOPUS database. The most of research papers are interdisciplinary nature out of a total of 2557 papers. Because of this interdisciplinary nature, papers have been classified in multiple subject areas and counted repeatedly in a total count of 3235 papers. A majority of papers which is 2121 (82.9 percent) related to Medicine followed by Biochemistry, Genetics and Molecular Biology (409), Pharmacology, Toxicology and Pharmaceutics (303), Neuroscience (86), Immunology and Microbiology (66), Environmental Science (51), Chemical Engineering (38), Social Sciences (29), Agricultural and Biological Sciences (27), Nursing (20), Health Professions (19), Dentistry (13), and Psychology (10). Less number of research works has been done in the field of computer science, economics and finance and decision science.

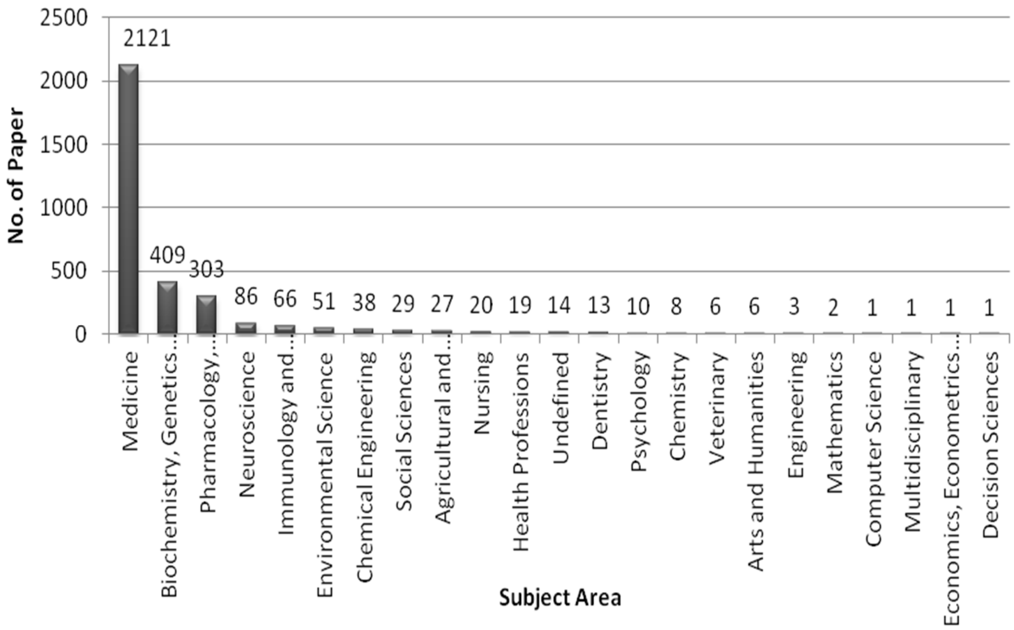


Figure 2: Subject-wise rank distribution of publication

Note: See table A in Appendix for complete list

Most prolific authors

There are 144 faculty and scientists are working in UCMS presently whose work has been covered in SCOPUS. Their contributions have been collected from the databases. They have contributed 2686 papers with 14356 citations at the rate of 5.34 citations per paper.

Table 2 reflects the name of top fifty most productive authors who have produced more than 20 papers. The top five authors have produced more than 67 papers by each. The first ranked author is B D Banaerjee (102 papers) followed by P. Gupta (Department of Paediatrics, UCMS) with 88 papers, which is 4.02 percent of total; the third ranks taken by A Singhal with 71 papers and fourth M S Bhatia with 68 papers. The analysis of authors based on citations they received for their papers as shown in table 2a, indicates that, Shukla, R. has got 17.92 average citations followed by Banerjee, B.D. with 16.16, Gambhir, J.K. (15.42), Sharma, S.B.(13.27), and Ahmed, R.S. with (12.98) citations per paper and all these five authors belongs to biochemistry department.

Table 2: Most prolific authors at the rank of fifteen

S N	Author	Department	Rank	Total Paper	Total Citation	Citation Per Paper
1	Banerjee, B.D.	Biochemistry	1	102	1648	**16.16**
2	Gupta, P.	Paediatrics	2	88	586	6.66
3	Singal, A.	Dermatology	3	71	241	3.39
4	Bhatia, M.S.	Psychiatry	4	68	450	6.62
5	Mediratta, P.K.	Pharmacology	5	67	572	8.54
6	Faridi, M.M.A.	Paediatrics	6	66	251	3.8
7	Jain, A.K.	Orthopaedics	7	64	595	9.3
8	Arora, V.K.	Pathology	8	61	382	6.26

9	Singh, N.	Pathology	9	58	391	6.74
10	Sharma, S.	Pathology	10	54	140	2.59
11	Pandhi, D.	Dermatology	11	52	147	2.83
12	Ahmed, R.S.	Biochemistry	12	50	649	**12.98**
13	Sharma, S.	Pathology	13	47	139	2.96
14	Tripathi, A.K.	Biochemistry	14	45	259	5.75
15	Aggarwal, A.	Paediatrics	15	44	141	3.2
16	Bhattacharya, S.N.	Dermatology	16	40	204	5.1
17	Madhu, S.V.	Medicine	17	39	312	8
18	Tyagi, A.	Anaesthesiology	17	39	198	5.08
19	Sethi, A.K.	Anaesthesiology	18	38	220	5.79
20	Chaturvedi, S.	Community Medicine	18	38	131	3.45
21	Sharma, S.B.	Biochemistry	19	37	491	**13.27**
22	Mohta, M.	Anaesthesiology	20	36	191	5.31
23	Gomber, S.	Paediatrics	21	35	297	8.48
24	Guleria, K.	Obstetrics & Gynaecology	22	33	163	4.94
25	Jain, B.K.	Surgery	23	32	217	6.78
26	Sunil Kumar	Surgery	23	32	95	2.97
27	Goel, N.	Obstetrics & Gynaecology	23	32	94	2.94
28	Rajaram, S.	Obstetrics & Gynaecology	23	32	82	2.56
29	Singh, U.R.	Pathology	24	31	106	3.42
30	Garg, P.K.	Surgery	25	30	10	0.33
31	Rusia, U.	Pathology	26	28	212	7.57
32	Shah, D.	Paediatrics	27	27	129	4.78
33	Singh, S.	Physiology	28	26	166	6.38
34	Gupta, A.	Surgery	28	26	111	4.27
35	Ramachandran, V.G.	Microbiology	29	25	210	8.4
36	Dewan, P.	Paediatrics	29	25	67	2.68
37	Gambhir, J.K.	Biochemistry	30	24	370	**15.42**
38	Sudhir Kumar	Orthopaedics	30	24	142	5.92
39	Kannan, A.T.	Community Medicine	30	24	60	2.5
40	Dhaliwal, U.	Ophthalmology	31	23	143	6.22
41	Agrawal, V.	Surgery	31	23	94	4.1
42	Vaney, N.	Physiology	31	23	82	3.56
43	Agarwal, M.P.	Medicine	31	23	41	1.78

44	Mohanty, D.	Surgery	31	23	31	1.35
45	Sikka, M.	Pathology	32	22	109	4.95
46	Kotru, M.	Pathology	32	22	48	2.18
47	Radhakrishnan, G.	Obstetrics & Gynaecology	33	20	102	5.1
48	Kalra, O.P.	Medicine	33	20	70	3.5
49	Suneja, A.	Obstetrics & Gynaecology	33	20	68	3.4
50	**Wadhwa, N.**	**Pathology**	**33**	**20**	**59**	**2.95**
144			52			
				2686	14356	5.34

Note: Number of results: 2557

Table 2a: Proliofic Authors based on Citations to their Papers

S N	Authors	Department	Rank	Total Papers	Total Citations	Citatiom/paper
1	Shukla, R.	Biochemistry	40	13	233	17.92
2	Banerjee, B.D.	Biochemistry	1	102	1648	**16.16**
3	Gambhir, J.K.	Biochemistry	30	24	370	**15.42**
4	Sharma, S.B.	Biochemistry	19	37	491	**13.27**
5	Ahmed, R.S.	Biochemistry	12	50	649	**12.98**
6	Singh, N.P.	Microbiology	37	16	174	10.87
7	Puri, D.	Biochemistry	41	12	128	10.67
8	Chhabra, P.	Community Medicine	37	16	162	10.12

Distribution of Authorship pattern

Table 3 shows a steady increase in publication of papers from 1975 to 2002 and a rapid progress during 2003 to 2012 with some exceptions. The collaborative coefficient is calculated 0.64 which denotes high degree of collaboration among authors. The authorship pattern shows that the highest contribution of 25.6 percent of total publications are made under triple authorship collaboration followed by four authorship (24.5%), double authorship (18. 6%), five authorship with 12.8% and so on. Most of the research works are team research.

Table 3: Distribution of Authorship pattern

AUTHORSHIP PATTERN Year	Single authored papers F1	Double authored papers F2	Three authored papers F3	Four authored papers F4	Five authored papers F5	Six authored papers F6	Seven authored papers F7	Eight authored papers F8	Nine authored papers F9	≥ Ten authored papers F10	Total N	ΣF	ΣF/N	CC
2013	8	30	28	42	24	21	11	4	0	3	171	53.5	0.31	0.69
2012	16	43	49	41	23	24	12	1	2	5	216	75.23	0.35	0.65
2011	12	24	50	73	32	13	6	3	3	10	226	70.02	0.31	0.69
2010	23	30	44	60	30	16	6	5	1	3	218	78.2	0.36	0.64
2009	19	24	43	34	19	10	6	1	0	3	159	60.56	0.38	0.62
2008	11	17	28	37	22	7	6	2	1	1	132	44.95	0.34	0.66
2007	11	19	24	26	15	5	0	0	0	1	101	38.93	0.38	0.61
2006	2	22	36	22	12	5	4	0	0	2	105	34.5	0.33	0.67
2005	4	25	32	27	13	5	0	0	0	1	107	37.44	0.35	0.65
2004	11	29	28	26	18	2	3	0	0	0	117	45.69	0.39	0.6
2003	7	17	31	30	19	4	0	0	0	2	110	37.99	0.35	0.65
2002	6	7	29	16	12	3	2	0	0	0	75	26.34	0.35	0.64
2001	5	16	18	16	9	2	1	0	0	1	68	25.37	0.37	0.63
2000	8	21	20	15	7	1	1	0	1	0	74	30.72	0.42	0.58
1999	9	22	18	9	11	3	1	0	0	0	73	31.09	0.43	0.57
1998	8	12	19	17	12	4	1	0	0	0	73	27.78	0.38	0.62
1997	4	15	29	12	3	3	1	0	0	1	68	25.5	0.37	0.63
1996	5	13	22	19	2	5	2	0	0	0	68	25.09	0.37	0.63
1995	5	8	15	12	6	3	0	0	0	0	49	18.7	0.38	0.62
1994	3	9	5	16	10	3	0	0	0	0	46	15.66	0.34	0.66
1993	3	14	13	13	6	3	0	0	0	0	52	19.28	0.37	0.63
1992	0	5	9	10	3	4	0	0	0	0	31	9.26	0.3	0.7
1991	6	14	14	9	3	1	0	0	0	0	47	20.67	0.44	0.56
1990	5	9	9	9	3	1	0	0	0	0	36	15.51	0.43	0.57
1989	0	6	6	4	3	0	0	0	0	0	19	6.6	0.35	0.65
1988	3	1	5	4	3	1	0	0	0	0	17	6.92	0.41	0.59
1987	0	2	4	3	3	0	0	0	0	0	12	3.68	0.31	0.69
1986	0	3	8	3	0	0	0	0	0	0	14	4.91	0.35	0.65
1985	0	2	2	5	0	1	0	0	0	0	10	4.07	0.41	0.59
1984	0	6	5	2	1	0	0	0	0	0	14	5.36	0.38	0.62
1983	1	3	2	3	0	1	0	0	0	0	10	4.07	0.41	0.59
1982	0	0	3	2	0	0	0	0	0	0	5	1.5	0.3	0.7
1981	0	3	0	3	2	0	0	0	0	0	8	2.65	0.33	0.67
1980	0	0	1	3	0	0	0	0	0	0	4	1.08	0.27	0.73
1979	1	1	2	2	0	0	0	0	0	0	6	2.66	0.44	0.56
1978	3	1	2	2	0	0	0	0	0	0	8	4.66	0.58	0.42
1977	2	1	0	0	1	0	0	0	0	0	4	2.7	0.67	0.33
1976	0	1	2	0	0	0	0	0	0	0	3	1.16	0.39	0.61
1975	1	0	0	0	0	0	0	0	0	0	1	1	1	0
Total	202	475	655	627	327	151	63	16	8	33	**2557**	921		
%	7.9	18.6	25.6	24.5	12.8	5.9	2.5	0.63	0.31	1.3	100			
Average													0.36	0.64

Note: Number of results: 2557,
CC: Collaborative Coefficient.

Degree of Collaboration

Table 4 reveals the degree of collaboration by calculating the pattern of single and joint authorship of papers. The degree of collaboration is 0.92 which means most of the research works are collaborative works. A number of 2355 papers out of 2557 papers are collaborative work.

Table 4: Degree of Collaboration

Year	Single authored papers (Ns)	Multi authored papers (Nm)	Nm + Ns	DC = $\frac{Nm}{Nm+Ns}$	Year	Single authored papers (Ns)	Multi authored papers (Nm)	Nm + Ns	DC= $\frac{Nm}{Nm+Ns}$
2013	8	163	171	0.95	1993	3	49	52	0.94
2012	16	200	216	0.92	1992	0	31	31	1
2011	12	214	226	0.95	1991	6	41	47	0.87
2010	23	195	218	0.89	1990	5	31	36	0.86
2009	19	140	159	0.88	1989	0	19	19	1
2008	11	121	132	0.92	1988	3	14	17	0.82
2007	11	90	101	0.89	1987	0	12	12	1
2006	2	103	105	0.98	1986	0	14	14	1
2005	4	103	107	0.96	1985	0	10	10	1
2004	11	106	117	0.9	1984	0	14	14	1
2003	7	103	110	0.94	1983	1	9	10	0.9
2002	6	69	75	0.92	1982	0	5	5	1
2001	5	63	68	0.93	1981	0	8	8	1
2000	8	66	74	0.89	1980	0	4	4	1
1999	9	64	73	0.88	1979	1	5	6	0.83
1998	8	65	73	0.85	1978	3	5	8	0.62
1997	4	64	68	0.94	1977	2	2	4	0.5
1996	5	63	68	0.93	1976	0	3	3	1
1995	5	44	49	0.9	1975	1	0	1	0
1994	3	43	46	0.93	**Total**	202	2355	2557	0.92

Note: Number of results: 2557,
DC: Degree of Collaboration.

Collaboration with other countries

The research collaboration among UCMS' researchers and international community is analyzed and 37 countries are ranked according to their degree of collaboration in table 5. Out of 2355 collaborative papers, 151 papers are written in collaboration with foreign authors, which covers 6.41 percent of total collaborative work. The majority of collaborative works (72.23 percent) are performed with eight countries (the United States, the United Kingdom, Nepal, Canada, Italy, Australia, Malaysia and Switzerland). USA is at rank 1st with 28.47 percent, subsequently UK (14.57 percent) at second, Nepal (12.58 percent) at third and Canada (6.0 percent) at fourth place. Italy, Australia, Malaysia and Switzerland are ranked fifth with 2.65 percent each and the foreign collaboration is gradually increasing.

Table 5: Country-wise distribution of collaboration

S N	Rank	Country	No. of publication	Percentage
1	1	United States	43	28.48
2	2	United Kingdom	22	14.57
3	3	Nepal	19	12.58
4	4	Canada	9	6.00
5	5	Italy	4	2.65
6	5	Australia	4	2.65
7	5	Malaysia	4	2.65
8	5	Switzerland	4	2.65
9	6	Germany	3	1.99
10	6	Iran	3	1.99
11	7	Slovakia	2	1.32
12	7	China	2	1.32
13	7	Fiji	2	1.32
14	7	Poland	2	1.32
15	7	France	2	1.32
16	7	South Korea	2	1.32
17	7	Sweden	2	1.32
18	7	Thailand	2	1.32
19	7	Uganda	2	1.32
20	8	Spain	1	0.66
21	8	Netherlands	1	0.66
22	8	Botswana	1	0.66
23	8	Rwanda	1	0.66
24	8	Saudi Arabia	1	0.66
25	8	Singapore	1	0.66
26	8	Brazil	1	0.66
27	8	South Africa	1	0.66
28	8	Cameroon	1	0.66
29	8	Austria	1	0.66
30	8	Finland	1	0.66
31	8	Guatemala	1	0.66
32	8	Hungary	1	0.66
33	8	Turkey	1	0.66
34	8	Kenya	1	0.66
35	8	Kuwait	1	0.66
36	8	Argentina	1	0.66
37	8	Zambia	1	0.66
		Total	151	100

List of core Journals

Core journals are identified and presented with their impact factor (2012) which is preferred by researchers for publishing their research work. A total of 1927 papers are published in 160 journals. The most prominent 20 journals are listed in table 6 according to their rank, which is calculated on the basis of occurrence of papers publications of UCMS' researchers in these journals. Indian Pediatics has published 187 papers and got first rank followed by Indian Journal of Physiology and Pharmacology at second rank with 107 papers, Indian Journal of Pediatrics in third with 77 papers and so on. In this list of 20 titles, 17 journals are published in India, which means Indian journals are mostly preferred by researchers for publication purpose.

Table 6: Most preferred journals with their impact factors

SN	Rank	Source Title	Country	No. of Papers	IF (2012)*
1	1	Indian Pediatrics	India	187	1.06
2	2	Indian Journal of Physiology and Pharmacology	India	107	0.65
3	3	Indian Journal of Pediatrics	India	77	0.86
4	4	Indian Journal of Medical Research	India	53	2.19
5	5	Journal of the Indian Medical Association	India	52	0.20
6	6	Indian Journal of Pathology and Microbiology	India	48	0.84
7	7	Indian Journal of Experimental Biology	India	47	1.59
8	8	Indian Journal of Dermatology Venereology and Leprology	India	46	1.206
9	9	Acta Cytologica	USA	44	0.94
10	10	Indian Journal of Orthopaedics	India	41	0.90
11	11	Journal of Anaesthesiology Clinical Pharmacology	India	40	0.47
12	12	Indian Journal of Medical Sciences	India	38	0.11
13	13	Journal International Medical Sciences Academy	India	37	0.03
14	14	Indian Journal of Clinical Biochemistry	India	33	1.27
15	15	Journal of Association of Physicians of India	India	29	0.53
16	16	Journal of Communicable Diseases	India	28	0.03
17	17	Tropical Doctor	UK	27	0.83
18	18	Indian Journal of Radiology and Imaging	India	25	0.56
19	19	Diagnostic Cytopathology	USA	23	1.22
20	19	National Medical Journal of India	India	23	0.71
21	20	Indian Journal of Medical Microbiology	India	22	1.19

* SCImago. (2007). SJR - SCImago Journal & Country Rank.

Major Findings

- A total of 2557 papers were published from 1975 to 2013 by the faculty members and researchers of UCMS. There is continuous increase in research output, which presents three types of growth in research publications, i.e. low growth, medium growth and high growth. In last six years (from 2008 to 2013) 300 papers are added in the research output which indicates fast increase and highly productive period.
- It is found that most of research papers are interdisciplinary nature out of a total of 2557 papers.

- A majority of papers which is 2121 (82.9 percent) related to Medicine followed by Biochemistry, Genetics and Molecular Biology (409), Pharmacology, Toxicology and Pharmaceutics (303), Neuroscience (86), Immunology and Microbiology (66), Environmental Science (51)
- The researchers are preferred Indian journals as core medium to publish their research output in medical sciences. Indian Pediatrics and Indian Journal of Physiology and Pharmacology were found to most preferred journals
- The authorship pattern illustrate that the highest contribution of 25.6 percent is made under three authors partnership.
- Overall degree of collaboration and average collaborative coefficient are found 0.92 and 0.64 respectively which indicate high trend of collaborative research among researchers.
- The first ranked author is B D Banaerjee (102 papers) followed by P. Gupta (Department of Paediatrics, UCMS) with 88 papers, which is 4.02 percent of total; the third rank by A Singhal with 71 papers. In terms of citation rate Shukla, R. has got 17.92 average citations followed by Banerjee, B.D. with 16.16, Gambhir, J.K. (15.42),
- The researchers are interested in collaborative research and the collaboration is made in various ways such as inter departments and inter institutes. They also made research cooperation with international community and 37 countries are identified with which collaborative researches have been conducted. USA is the most preferred country, subsequently UK (14.57 percent) at second and Nepal (12.58 percent) at third by the researchers for research cooperation.

References

British Standards Institution. (1976). *British standards of documentation terms.* London: BSI.

Dhawan, S. M. & Gupta, B. M. (2007). Physics Research in India: A Study of Institutional Performance based on Publications Output. *DESIDOC Bulletin of Information Technology,* 27(1), 55-67.

Kao, C. & Pao, H. L. (2009). An evaluation of research performance in management of 168 Taiwan universities. *Scientometrics,* 78(2), 261–277.

Kumar, H. A. & Dora, M. (2012). Research Productivity in a Management Institute: An Analysis of Research Performance of Indian Institute of Management Ahmedabad during 1999-2010. *DESIDOC Journal of Library & Information Technology*, 32(4), 365-372.

Maharana, R. K. & Sethi, B. B. (2013). A bibliometric analysis of the research output of Sambalpur University's publication in ISI Web of Science during 2007-11. *Library Philosophy and Practice (e-journal).* Retrieved from http://digitalcommons.unl.edu/libphilprac/926

Ministry of Health and Family Welfare, Government of India. (2012). *National Health Profile (NHP) of India, 2012.* Retrieved from http://cbhidghs.nic.in/index2.asp?slid=1256&sublinkid=1163

Ministry of Science and Technology, Government of India. (2013). *Science, Technology and Innovation Policy 2013.* Retrieved from http://www.dst.gov.in/sti-policy-eng.pdf

Prathap, G. & Gupta, B. M. (2009). Ranking of Indian universities for their research output and quality using a new performance index. *Current Science*, 97(6), 751-752.

Savanur, K. & Konnur, P. V. (2012). Growth and impact of research output of Bangalore University, 1971-2010: A scientometric study. *International Journal of Library and Information Scie*nce, 3(5), 71-80.

Sevukan, R. & Sharma, J. (2008). Bibliometric analysis of research output of biotechnology faculties in some Indian central universities. *DESIDOC Journal of Library & Information Technology,* 28(6), 11-20.

Thomson Reuters (2012). *Bibliometric study of India's Scientific Publication outputs during 2001-10: Evidence for changing trends.* Retrieved from http://www.dst.gov.in/whats_new/whats_new12/report.pdf

UCMS (2013). Retrieved from: http://www.ucms.ac.in

Vasishta, S. (2011). Assessment of Academic Research Output during 1996-2009: A Case Study of PEC University of Technology, Chandigarh. *DESIDOC Journal of Library & Information Technology*, 31(2), 136-142.

Gupta, B. M., Bala, A. & Kshitig, A. (2013). World Cataract Research: A Scientometric Analysis of Publications Output during 2002-11. *Library Philosophy and Practice (e-journal)*. Retrieved from: http://digitalcommons.unl.edu/libphilprac/895

SCImago. (2007). SJR- SCImago Journal & Country Rank. Retrieved May 19, 2014, from http://www.scimagojr.com

Relationship between Development Indicators and Contribution to the Science: Experiences from Iran

Mohammad Hassanzadeh* and Babak Akhgar**

*Dean, Knowledge and Information Science Department, Tarbiat Modares University
hasanzadeh@modares.ac.ir

**Sheffield Halam Univarsity
baakhgar@yahoo.com

Abstract

This paper aimed in investigation of the relationship between selected development indicators and contribution to the global knowledge in provincial level. The research was focused on Iranian papers published in Thompson Reuters in 2012. Data was collected running an advanced search limiting the results to the papers published with their address belong to the Islamic republic of Iran. Nearly, 28.000 scientific records were analyzed. Then we run applications to separate records based on the affiliation province of the authors. Contribution of each province analyzed against selected developmental indicators such as human development indicator (HDI), family size, healthcare centers, population, universities and students. Findings showed that there is significant relationship between science infrastructure and developmental indicators and also their contribution to the knowledge. But amongst them relationship between family size and science production was negative. Based on findings we can conclude that scientific progress has its grassroots in developmental programs and in turn, developmental achievements will flourish the science in societies.

Keywords: Science indicators, development indicators, Iran, Provinces, relationship, Policy making

Introduction

In the last Collnet conference in Stoni, we proposed "Kientometrics" (Hassanzadeh, Akhgar, 2013) instead of scientometrics. Our main reason for this naming was that this approach will improve the interoperation between science indicators and knowledge society indicators and will provide science policy makers with good understanding of impacts of science in knowledge society and also implications of the knowledge society on scientific progresses. With reference to that notion we believe that knowledge gradually turns into the most important enabler especially for developing countries solving their problems and improving public welfare and act as a remedy for their old pain. Expenditures on science and technology experience an increasing trend despite even economic downturn and income difficulties almost in all countries. This indicates that, governments and public see the science and technology as rescuing Savior angel for their societies. But in other hand, under development and subsequent barriers work as prohibiting and deterrent factor against the growth of the science. Therefore, growth of science in developing countries always remains ambiguous. This hazy situation results in decreased public support of science. Finally, potential financial resources for scientific research is limited only to governmental ever decreasing budgets and this destructive cycle perpetuates and complicates the scientific growth in this countries.

The relationship between development indicators and science growth works also in local governments and provincial level. Less developed local states encounter with low speed of science growth and low participation in national scientific progress. This level of analysis often ignored in research initiatives and nation- wide investigations.

It seems that, there is significant relationship between development indicators and scientific progress in societies especially in developing countries. This convinces us to take into account the development indicators in any initiative of analyzing scientific progress. Development indicators like human resource development, life expectancy, education and income affect the participation of provinces in national science.

This paper strived to articulate the relationship between development indicators and scientific progress and recommend a framework for future nation-wide and international investigations on scientific productions. First Section of this investigation belongs to factual data. In this section, participation of each province in total national science production and its situation on development indicators has been described. The Second part of the paper allocated to the relationship between development indicators and science growth. Finally, the Third section discusses the implications of findings on scientometrics methodologies and analytics and some recommendations proposed at this respect.

Methodology

Data was collected from domestic and international citation databases ISC[1] (Islamic Science Citation Center) and Thomson Reuters (former ISI[2]) for Iran. We carried out an advance search in these data bases for Iranian papers published in 2012. After downloading then in plain text format classified them by provinces[3]. In another hand we extracted development indicators for 31 provinces from Iran national agency for statistics[4]. We selected literacy rate, family size, urban population and human development index (HDI) as developmental indicators. Furthermore, we collected data about some contextual factors accelerating scientific researches such as universities and research centers, students and population growth in provinces.

To examine relationship between development indicators and science indicators, we carried out a Pearson correlation using Spss 19.0 and for more visualization of the correlation pattern we have depicted it by scatter plot.

Findings

The results of the study indicate overall significant relationship between developmental indicators and their participation rate in national scientific productions. Less developed states have low participation and vice versa. But there are some intervening factors which should be taken into account in any policy making decision. Factors such as student mobility, visiting professors and interprovincial co operations play crucial role and should be cleared for policy purposes.

Table 1. Correlation coefficient in Intera-indicator level: Developmental indicators

		Healthcare Centers	Urbanization rate	Literacy rate	family size
HDI	Pearson Correlation	.314	.542(**)	.691(**)	-.457(**)
	Sig. (2-tailed)	.085	.002	.000	.010
	N	31	31	31	31
Healthcare Centers	Pearson Correlation	1	.094	.132	-.182
	Sig. (2-tailed)	.	.615	.478	.326
	N	31	31	31	31
Urbanization rate	Pearson Correlation	.094	1	.536(**)	-.389(*)

[1] www.isc.gov.ir
[2] www.isiknowledge.com
[3] Iran has 31 provinces
[4] www.amar.org.ir

	Sig. (2-tailed)	.615	.	.002	.031
	N	31	31	31	31
Literacy rate	Pearson Correlation	.132	.536(**)	1	-.489(**)
	Sig. (2-tailed)	.478	.002	.	.005
	N	31	31	31	31
family size	Pearson Correlation	-.182	-.389(*)	-.489(**)	1
	Sig. (2-tailed)	.326	.031	.005	.
	N	31	31	31	31
population growth rate	Pearson Correlation	.062	.155	.476(**)	.066
	Sig. (2-tailed)	.740	.404	.007	.724
	N	31	31	31	31

** Correlation is significant at the 0.01 level (2-tailed).
* Correlation is significant at the 0.05 level (2-tailed).

In the intera indicator level, there was significant relationship between HDI and three other indicators (Urbanization rate, Literacy rate and family size). Relationship between the HDI and family size was negative. It means that, any increase in family size coincide with decrease in human development index value. This is mainly because family size directly affects the economic tempore of the family which is the main stimuli for education and development. This in turn, decreases the ability of the people to continue education and as well as research affaires which results in publishing scientific papers. In another hand there was negative relationship between family size and urbanization and also between literacy rate and family size. These results urge policy makers to include interrelationship between developmental indicators in any social and scientific policy formulation. The next table summarizes the relationship among scientific – infrastructural indicators.

Table 2. Correlation coefficient in Intera -indicator level: Science-infrastructure indicators

		Population share	Academic centers share	University student share	Science per population	Science per student
Science production	Pearson Correlation	.769(**)	.577(**)	.879(**)	.739(**)	.842(**)
	Sig. (2-tailed)	.000	.001	.000	.000	.000
	N	31	31	31	31	31
Population share	Pearson Correlation	1	.561(**)	.851(**)	.220	.465(**)
	Sig. (2-tailed)	.	.001	.000	.234	.008
	N	31	31	31	31	31
Academic centers share	Pearson Correlation	.561(**)	1	.616(**)	.424(*)	.481(**)
	Sig. (2-tailed)	.001	.	.000	.018	.006
	N	31	31	31	31	31
University student share	Pearson Correlation	.851(**)	.616(**)	1	.488(**)	.517(**)
	Sig. (2-tailed)	.000	.000	.	.005	.003
	N	31	31	31	31	31
Science per population	Pearson Correlation	.220	.424(*)	.488(**)	1	.861(**)
	Sig. (2-tailed)	.234	.018	.005	.	.000
	N	31	31	31	31	31
Science per student	Pearson Correlation	.465(**)	.481(**)	.517(**)	.861(**)	1
	Sig. (2-tailed)	.008	.006	.003	.000	.
	N	31	31	31	31	31
Science per university	Pearson Correlation	.568(**)	.099	.669(**)	.662(**)	.762(**)

	Sig. (2-tailed)	.001	.595	.000	.000	.000
	N	31	31	31	31	31

** Correlation is significant at the 0.01 level (2-tailed).
* Correlation is significant at the 0.05 level (2-tailed).

Results showed that there is significant relationship between science production and other science demographic indicators such as population (r=0.769), students (r=0.879) and also universities (r=0.577). The more the number of universities in a province, the more their share in national science production. And naturally, increase in the number of universities in a province coincide with the increase in the number of students who carry out scientific investigations and finally publish their findings in scientific journals. Academic centers share of provinces and also number of students indicated significant relationship with other indicators. As the science infrastructures in a province grows, consequently its ability and opportunity to produce scientific outputs also increase. Scientific progress is a process which has its roots in human resources who has access to the academic opportunities and facilities to carry out research initiatives and publish scientific outputs. Any policy formulation and also science evaluation initiatives in provincial level should take in account the academic infrastructure and human development indicators to reach in valid conclusions.

Table 3. Correlation coefficient in between science and development

		HDI	Healthcare Centers	Urbanization rate	Literacy rate	family size
Science production	Pearson Correlation	.363(*)	.730(**)	.392(*)	.356(*)	-.465(**)
	Sig. (2-tailed)	.045	.000	.029	.050	.008
	N	31	31	31	31	31
HDI	Pearson Correlation	1	.314	.542(**)	.691(**)	-.457(**)
	Sig. (2-tailed)	.	.085	.002	.000	.010
	N	31	31	31	31	31
Healthcare Centers	Pearson Correlation	.314	1	.094	.132	-.182
	Sig. (2-tailed)	.085	.	.615	.478	.326
	N	31	31	31	31	31
Urbanization rate	Pearson Correlation	.542(**)	.094	1	.536(**)	-.389(*)
	Sig. (2-tailed)	.002	.615	.	.002	.031
	N	31	31	31	31	31
Literacy rate	Pearson Correlation	.691(**)	.132	.536(**)	1	-.489(**)
	Sig. (2-tailed)	.000	.478	.002	.	.005
	N	31	31	31	31	31
family size	Pearson Correlation	-.457(**)	-.182	-.389(*)	-.489(**)	1
	Sig. (2-tailed)	.010	.326	.031	.005	.
	N	31	31	31	31	31
population growth rate	Pearson Correlation	.273	.062	.155	.476(**)	.066
	Sig. (2-tailed)	.138	.740	.404	.007	.724
	N	31	31	31	31	31

* Correlation is significant at the 0.05 level (2-tailed).
** Correlation is significant at the 0.01 level (2-tailed).

Relationship between contribution to the knowledge and selected developmental indicators is significant. Amongst them the relationship between science production and family size is negative and significant. As we referred to in table 1, the relationship between family size and HDI also was negative. This indicates that dealing with science affaires in provincial level may be affected by economical situation of families and their ability to provide their members with good education opportunity and facilities. In the other hand any evaluation of science in provincial level should consider the developmental indicators and economical factors which affect the ability of families to contribute in knowledge initiatives either positive or negative. Rationality for this can be found in table 3, where Family size has negative relationship with literacy rate. This means that as family size grows, conversely, the literacy rate of people decreases. The interesting finding was that relationship between the number of healthcare centers in provinces and their contribution in global and national knowledge was significant and the highest among others. The number of health care centers in provinces by itself do not seem to be an accelerator of scientific affairs but the highest relationship between this variable and science production urge us more acceptance of the Iranian proverb that “the wisdom will be accommodate in the mind of the healthy man”.

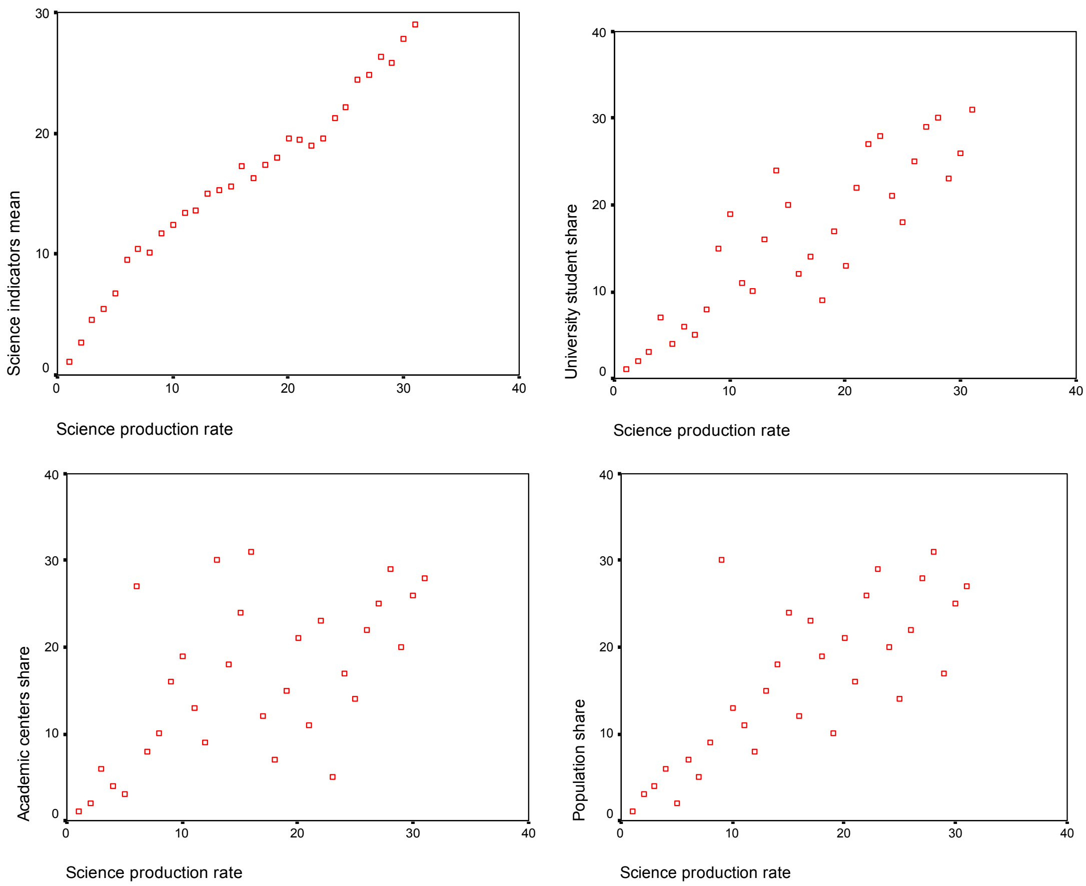


Figure 1. Scatter plot for science-infrastructure variables

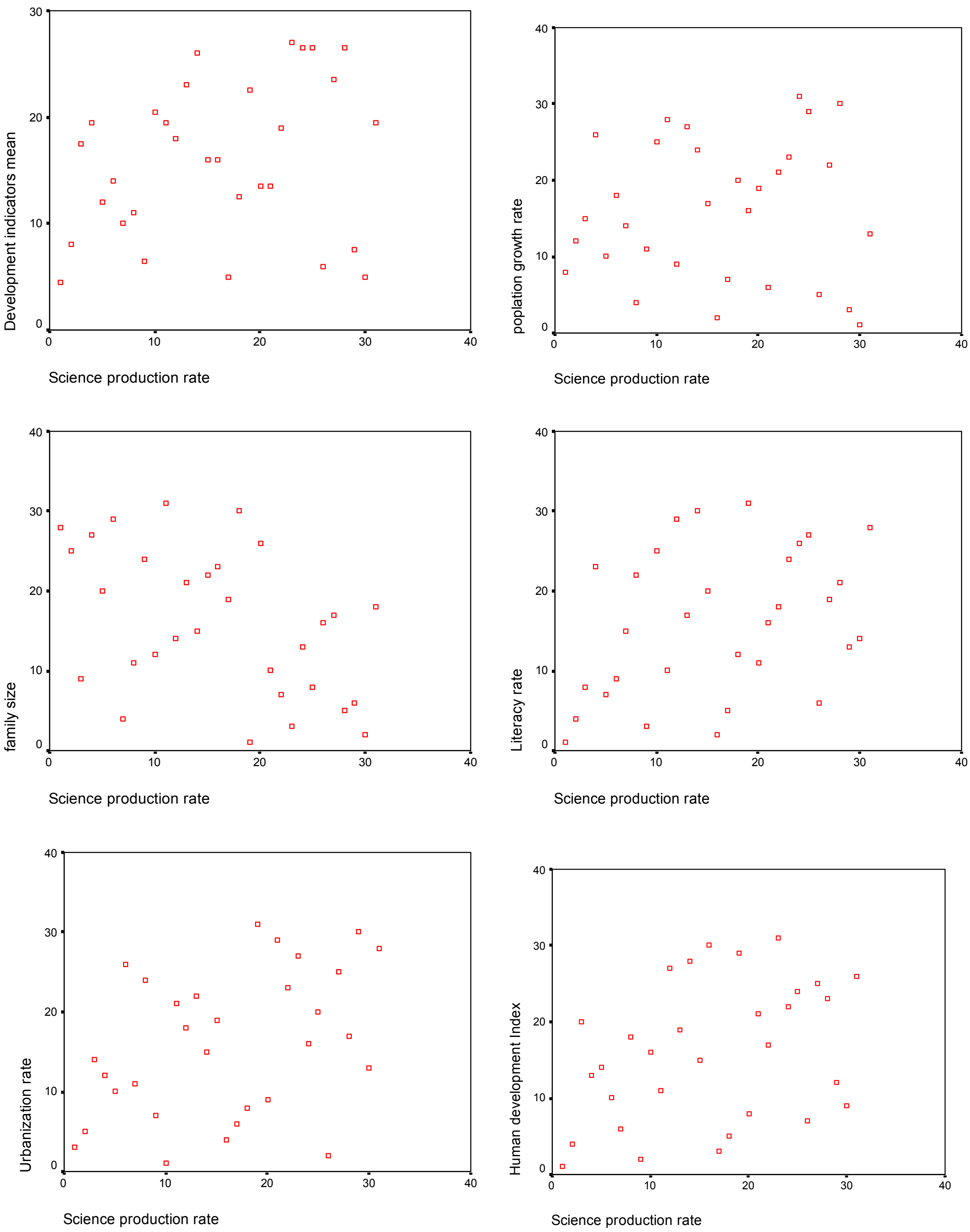


Figure 2. Scatter plot for science –developmental variables

Furthermore, infrastructures enabling researchers to conduct sophisticated research projects and collaborations with national and international academic centers will affect the scientific productivity of the provinces should be counted. Transition of a society from underdeveloped to developing and from this point to developed one is a complex process that requires suitable synchronization of important actors such as scientists and civil servants. In other word, the impetus for this transformation has to come not only from scientists but from other social actors as well. In a world where poverty, inequality and many social and environmental challenges go to grow and knowledge-based economy offer breakthrough solutions, sustainable progress requires that policy makers coincide scientific and technological capabilities and developmental goals and formulate effective strategies for the future harmonic developments through scientific progress.

Conclusion remarks

Findings of this research will contribute scientometrics community with insight which has come from real world. Application of Such an approach to scientometrics studies not only will improve the effectiveness of this kind of investigations, but also will attract attentions from executive and decision making community to them as well. Integrating science indicators with developmental indicators will help science policy makers to formulate right strategies to improve synergic nature of science and development and provide societies with good alignment of the science strategies and development goals.

We conclude that scientific progress is a complex process which contains a multidimensional framework. In this framework, the role of human resources, Research and development facilities are outstanding. But this is not the all of the story, developmental initiatives and related indicators play crucial role in flourishing scientific systems. Significant relationship between developmental indicators and contribution into knowledge in provincial level urge us to revise our science policies from focus on R&D to science and developmental goals.

Almost all of science and technology indicators heavily focused on S&T expenditures and S&T personnel. Meanwhile, we should take into consideration the deeper layers of the science and technology: Developmental goals. This approach will guide us to consider main social and real world problems related with science and technology. This requires major revisions in current approach to the science and technology evaluation and more reflections on the indicators and tools we are using to measure the science. We should not forget that the ultimate goal of the science is creating the most prosperous future for human being. For that, it is better to bridge the existing gap in our framework and tools by integrating knowledge, science and developmental indicators.

References

Hassanzadeh, Mohammad; Akhgar, Babak and Navidi, Fatemeh (2013). Integrating Knowledge and Science indicators (Kientometrics) towards a rational framework of investigations: a Comparative Approach. *9th international conferences on Webometrics, Informetrics and Scientometrics & fourteenth COLLNET Meeting*. September 3-5, 2013 Stoni.

European cancer research publications, 2002-13

Mursheda Begum and Grant Lewison

King's College London, Research Oncology, Guy's Hospital, Great Maze Pond, London SE1 6RT, UK
mursheda.begum@kcl.ac.uk, grantlewison@aol.co.uk

Abstract

This study was undertaken as part of an EU research mapping exercise concerned with five non-communicable diseases (NCDs) – cancer, cardiovascular disease, diabetes, mental disorders and respiratory diseases. It was designed to investigate the funding of European NCD research, any gaps or overlaps, and its influence on clinical guidelines, stories in the mass media and governmental health policy statements. (A poster at this conference describes the analysis of the papers cited by cancer clinical guidelines). Europe was defined as the 28 Member States plus Iceland, Norway and Switzerland, and the period covered was the 12 years from 2002 to 2013. In this paper, we describe the research outputs from the 31 countries, their international collaboration, research levels and citation scores, and their partition by cancer manifestation (site) and research type (e.g., chemotherapy, genetics, surgery). Comparisons with the countries' cancer disease burden suggest ways in which their cancer research portfolios could be made more responsive to need.

Introduction

This paper reports on the first stage of a two-year project supported by the European Union that is designed to map European research outputs in five non-communicable diseases and investigate their funding and their impact. Outputs are to be mapped in cancer (reported here), cardiovascular disease (including stroke), diabetes, mental disorders and respiratory diseases. The project involves seven partners in six EU member states: Estonia, France, Germany, Italy, Spain and the UK; it is managed and co-ordinated by the London School of Economics. King's College London is responsible for the bibliometrics, including the impacts of the research publications.

There does not appear to be any comprehensive study of cancer publications in the EU, although there are bibliometric studies of different samples of the total *oeuvre*. Thus there are articles on the output of papers on particular cancers, such as bladder (Kunath *et al*, 2013), breast (Glynn *et al*, 2010; Healy *et al*., 2011; Brolmann *et al*., 2012; Perez-Santos & Anaya-Ruiz, 2013), laryngeal (Glynn *et al.*, 2012) and lung (Ho *et al*., 2013). There are also studies on individual countries, such as Italy (Ugolini *et al.*, 1997), Norway (Skovlund, 1998), the UK (Lewison, 2003) and Canada (Campbell *et al*., 2010). Finally, some authors have examined the outputs of different research types, including clinical trials (Grossi *et al*., 2003), cancer nursing (Zhang *et al*., 2011) and quality-of-life issues (Bailey *et al*., 2010). In this paper, we describe the results of a study that encompasses all three categories – cancer sites, countries and research types – in the 28 Member States of the EU plus Iceland, Norway and Switzerland from the 12 years, 2002-13.

Methodology

Cancer research was defined by means of a complex filter originally devised in consultation with Cancer Research UK and recently updated with the aid of our Spanish partners in the project, the Escuela Andaluza de Salud Publica, S.A.. It consisted of long lists of specialist oncology journals and title words, including the various types of cancer, genes that increase individuals' chance of having particular cancers, and drugs used exclusively for the treatment of cancer. The filter was modified several times in order to improve its precision, p, and recall, r, and the final version had $p = 0.95$ and $r = 0.98$.

It was applied to the Web of Science (WoS) and articles and reviews in 31 European countries, see Table 1, were identified and downloaded to files from both the SCI and the SSCI. Five-year citation scores were also obtained for the papers from 2002-09. The details of the papers were transferred to an Excel spreadsheet by means of special macros for analysis, and the downloaded citation files were transformed by another macro so that the paper citation scores could be calculated and then transferred to the papers in the original spreadsheet, which contained details of 282,055 papers.

Table 1. List of 31 countries used to limit the ONCOL papers whose details were obtained.

ISO	*Country*	*ISO*	*Country*	*ISO*	*Country*	*ISO*	*Country*
AT	Austria	EE	Estonia	IS	Iceland	PL	Poland
BE	Belgium	ES	Spain	IT	Italy	PT	Portugal
BG	Bulgaria	FI	Finland	LT	Lithuania	RO	Romania
CH	Switzerland	FR	France	LU	Luxembourg	SE	Sweden
CY	Cyprus	GR	Greece	LV	Latvia	SI	Slovenia
CZ	Czech Rep.	HR	Croatia	MT	Malta	SK	Slovakia
DE	Germany	HU	Hungary	NL	Netherlands	UK	United Kingdom
DK	Denmark	IE	Ireland	NO	Norway		

The spreadsheet was annotated with 31 additional columns each of which contained the product of the paper's citation score, ACI, with the fractional presence of each country among its addresses. The sum of these products, divided by the fractional count of the country for the relevant years (in the first instance, the eight years 2002-09), then gave the country's citation score on a fractional count basis, which is more appropriate than the score based on integer counts. These individual country scores could then be compared with the ACI values for the world. These were obtained for each year's ONCOL publications directly from the WoS, although the sets of papers needed to be divided into sub-sets, based on journal initial letters, in order that each one should have no more than 10,000 papers, as this is the limit in the WoS for citation reports.

Then a succession of macros provided for each paper fractional counts of countries from the addresses[1], the research level (1=clinical, 4=basic) based on words in the title (Lewison & Paraje, 2004), the type of research (*e.g.*, chemotherapy, genetics, surgery) and the cancer site (one of 22 selected sites, *e.g.*, breast, lung, prostate)[2]. The identification of the research type(s) and cancer site(s) for each paper was performed by two further macros, each based on sub-filters created in consultation with Professor Richard Sullivan of KCL that consisted of title word strings and (for many of them) journal name strings. Table 2 lists the research types, with four-letter (tetragraph) codes and Table 3 the cancer manifestations, which corresponded closely to the ones listed in the recently-published world disease burden estimates (Murray *et al.*, 2012).

[1] For example a paper with two French addresses and one from Germany would be classified as FR = 0.67, DE = 0.33.

[2] Not all papers could be linked to a cancer site or to a research type, but some had more than one.

Table 2. List of research types in cancer research defined by sub-filters.

Research type	*Code*	*Research type*	*Code*	*Research type*	*Code*
Chemotherapy	CHEM	Palliative care	PALL	Radiotherapy	RADI
Diagnosis	DIAG	Pathology	PATH	Screening	SCRE
Epidemiology	EPID	Prognosis	PROG	Surgery	SURG
Genetics	GENE	Quality of life	QUAL		

Table 3. List of 22 cancer manifestations (body sites) for which sub-filters were developed to identify relevant ONCOL papers.

Site	*Code*	*Site*	*Code*	*Site*	*Code*
bladder	BLA	liver	LIV	pancreas	PAN
bone	BON	lung, trachea, bronchus	LUN	prostate	PRO
brain	BRA	lymphoma	LYM	stomach	STO
cervix	CER	breast	MAM	testicles	TES
colon / rectum	COL	melanoma	MEL	thyroid	THY
gallbladder	GAL	mouth (head & neck)	MOU	uterus	UTE
kidney	KID	oesophagus	OES		
leukaemia	LEU	ovaries	OVA		

A recent publication by the World Health Organization provides detailed estimates of the burden of disease (both deaths and Disability-Adjusted Life Years, DALYs) for each country and for many individual diseases for the year 2010. The data are provided both as different-sized rectangles within a square representing a country's (or region's, or the whole world's) total disease burden, and they can also be downloaded to file. We did this for the 31 countries of the European region, and for the disease areas relevant to this study; the data selected were for all ages and both sexes. They are in the form of percent of total DALYs for the country, and were then multiplied by the DALY total to give the DALYs for each disease and country. These could then be added to give the total for the EUR31 region, and the pattern of disease burden for each country compared with the European average. For some diseases, the differences were not great, but for others there were big variations in relative burden between countries. For cancer, data were provided on some 24 different manifestations, not all of which corresponded to our analysis of sites (see Table 3 above). However DALYs were provided for 13 sites whose details are given in this paper.

Results

Outputs of cancer research papers by European countries.

For each of the original 31 countries, we determined the integer and fractional count totals, and the numbers in each of the 12 years; we also detemined the annual average percentage growth rate (AAPG) based on fractional counts. [This was obtained from a plot of the logarithm of the number of papers each year.] Table 4 lists the results: since research output tends to be correlated with Gross National Product (rather than simply with population), we have plotted the countries' fractional paper counts against GDP for a representative year in Figure 1.

This table shows that there are big differences in output, with more than three orders of magnitude between the largest (Germany) and the smallest (Malta). However, some of the smaller countries are expanding their output rapidly – notably Romania, whose fractional count output rose from only 7 papers in 2002 to over 250 in 2013. The comparison with GDP suggests that some countries are publishing much more than their wealth would suggest, notably Iceland (x 2.8), Croatia (x 2.5), Slovenia (x 2.2) and Greece (x 2.0). On the other hand, some other countries are doing much less research than expected, such as Luxembourg (29%), Latvia (40%), Cyprus (53%) and France (62%).

Table 4. Outputs of 31 European countries in cancer research (ONCOL), 2002-13 (12 years) in both the SCI and SSCI. Integer (Int) and fractional (Frac) counts, % foreign contribution and the annual growth rate (aapg). ***Countries are ranked by fractional outputs. Codes: see Table 1.***

Country	*Int*	*Frac*	*% for*	*AAPG*		*Country*	*Int*	*Frac*	*% for*	*AAPG*
DE	60456	45436	24.8	2.6		IE	3367	2247	33.3	9.3
IT	48499	37876	21.9	4.8		PT	3136	2079	33.7	13.3
UK	52465	37541	28.4	2.4		HU	2855	1897	33.6	3.2
FR	40329	30127	25.3	4.1		HR	1720	1429	16.9	9.7
NL	23572	16068	31.8	4.5		RO	1748	1248	28.6	35.7
ES	21453	15654	27.0	7.6		SI	1298	898	30.8	10.6
SE	14881	9205	38.1	2.0		SK	1196	755	36.9	6.6
PL	9699	7543	22.2	10.0		BG	673	453	32.6	10.4
GR	9513	7243	23.9	3.8		LT	396	265	33.0	16.4
CH	12827	6837	46.7	4.1		IS	509	208	59.1	3.7
BE	10891	6253	42.6	2.9		LU	259	116	55.3	14.6
AT	8971	5563	38.0	1.1		EE	208	97	53.2	4.0
DK	7692	4713	38.7	8.0		LV	191	86	55.2	7.3
NO	6650	4054	39.0	6.2		CY	198	79	60.1	18.0
FI	6015	3721	38.1	0.0		MT	51	22	56.5	12.1
CZ	4422	3005	32.0	9.2						

It is also expected that researchers in the scientifically larger countries (e.g., UK, Germany) would find it easier to work with a partner within the country that provided complementary expertise than researchers from small countries (e.g., Estonia, Ireland) and would therefore tend to collaborate less internationally. However we might expect that international transnational links would be much weaker for the Member States in eastern Europe, and so Figure 2 has been plotted to show if this is the case. The figure shows that these "accession" Member States do indeed collaborate less than expected, whereas the five Scandinavian countries, with Belgium, Luxembourg and Switzerland, collaborate internationally more than the trend-line would suggest.

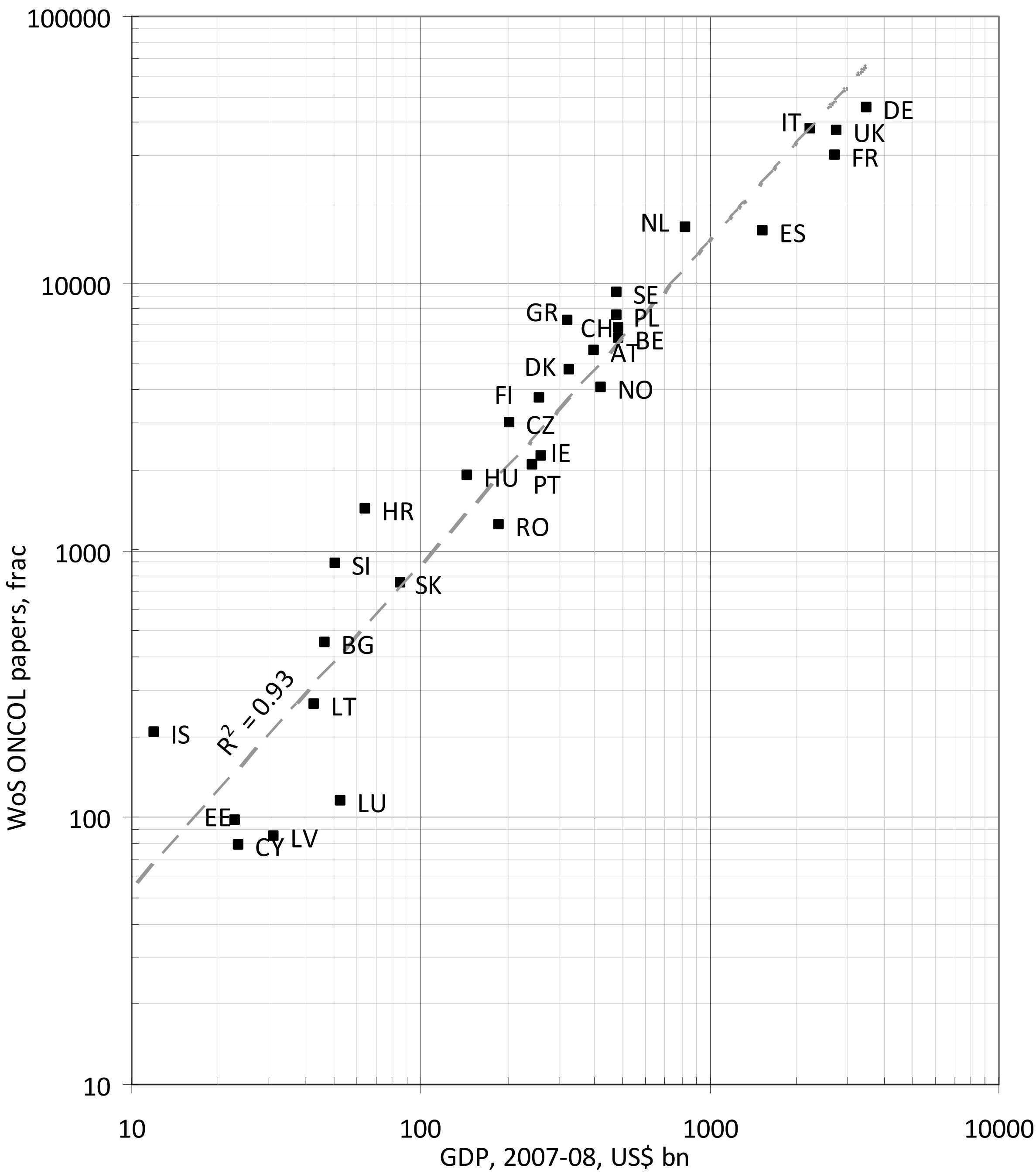


Figure 1. Plot of cancer research output, 2002-13, against GDP for European countries.

Note: MT omitted.

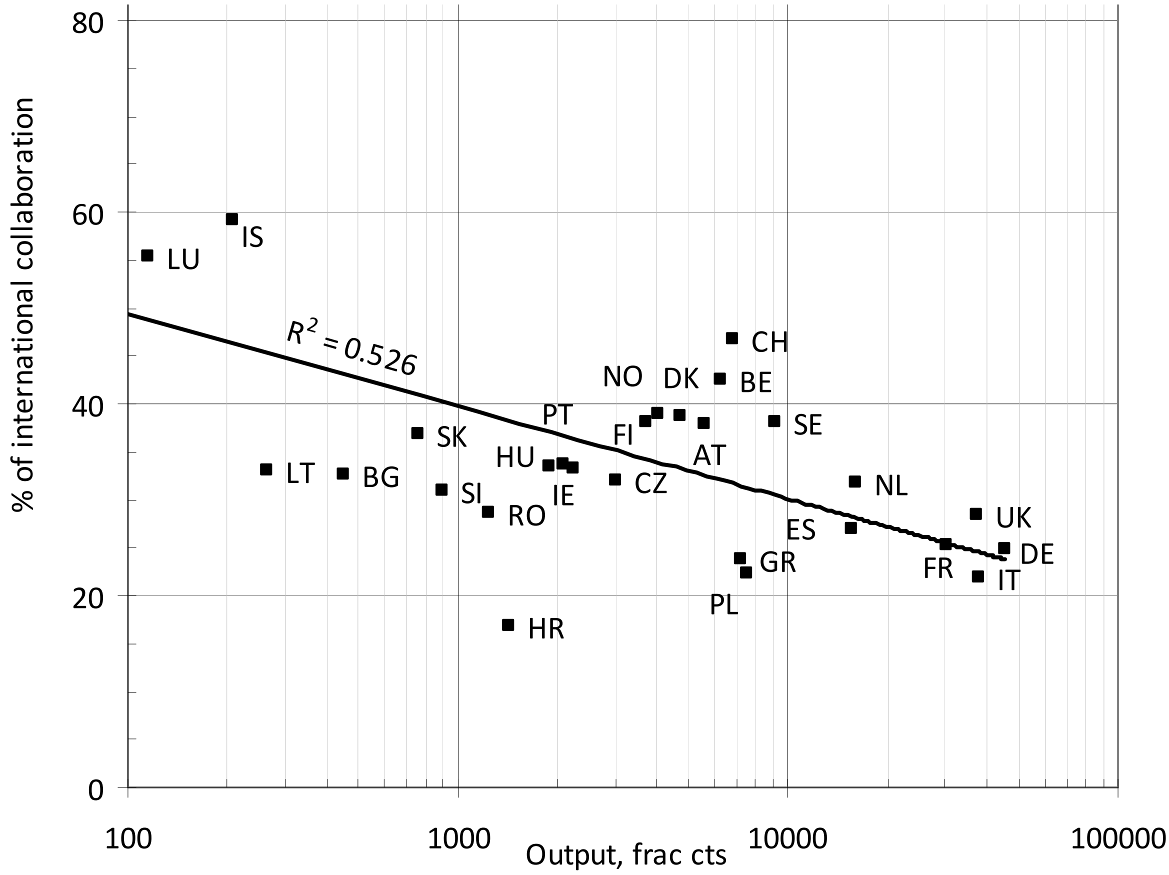


Figure 2. Percentages of international collaboration in cancer research, 2002-13, by European countries plotted against their output (fractional counts of papers).

Citation scores and percentage of reviews

Citation scores in most subject areas have been increasing slowly with time, in part because the WoS now covers more journals than previously, and also because authors are expected to be more punctilious in their acknowledgement of earlier work. Figure 3 shows the progression in cancer research ACI scores from 2002 to 2009; the values for intermediate years (2003-08) for Europe are shown as three-year moving averages in order to smooth out annual fluctuations. The mean score for Europe was slightly below the world average in 2002-03, but since 2006 it has been slightly higher, partly because of the greatly increased world presence of China, whose papers tend to be less well cited than average. Its papers have been becoming slightly more clinical (RL p), although the journals in which they have been published have altered little in terms of research level (RL j), Figure 4.

The mean citations per paper for the EUR31 countries are shown in Table 5. This also shows how many of a country's papers received enough cites to put them in the top 5% of EUR31 papers in the eight-year period, for which the qualification was 53 cites. [There were actually 5.15% of European papers that achieved this number of citations.] This may be a better measure of how effective a country's research output is because it is normally the most influential papers that are really important to the development of a field. The two indicators are positively correlated, with $r^2 = 0.94$.

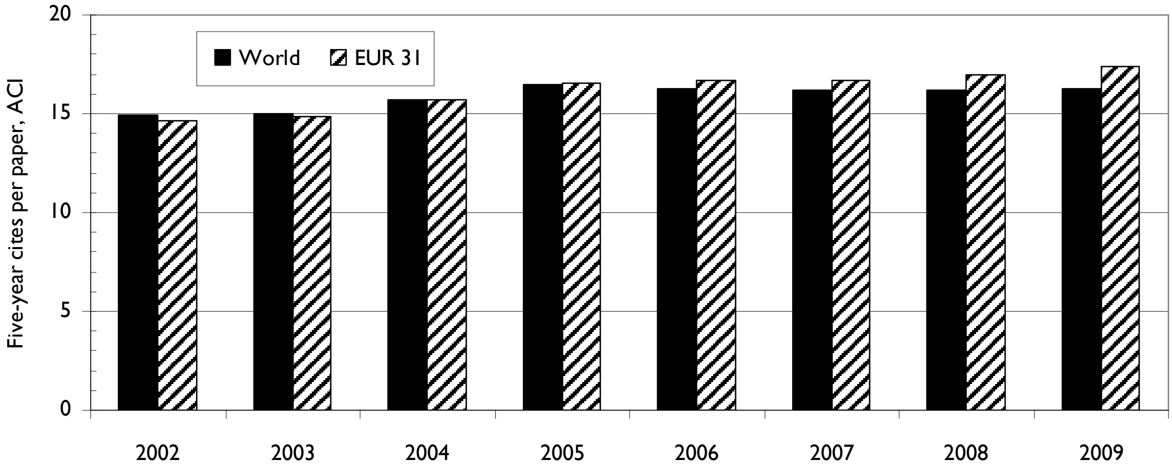


Figure 3. Mean values of five-year citation counts for world and EUR31 papers, 2002-09.

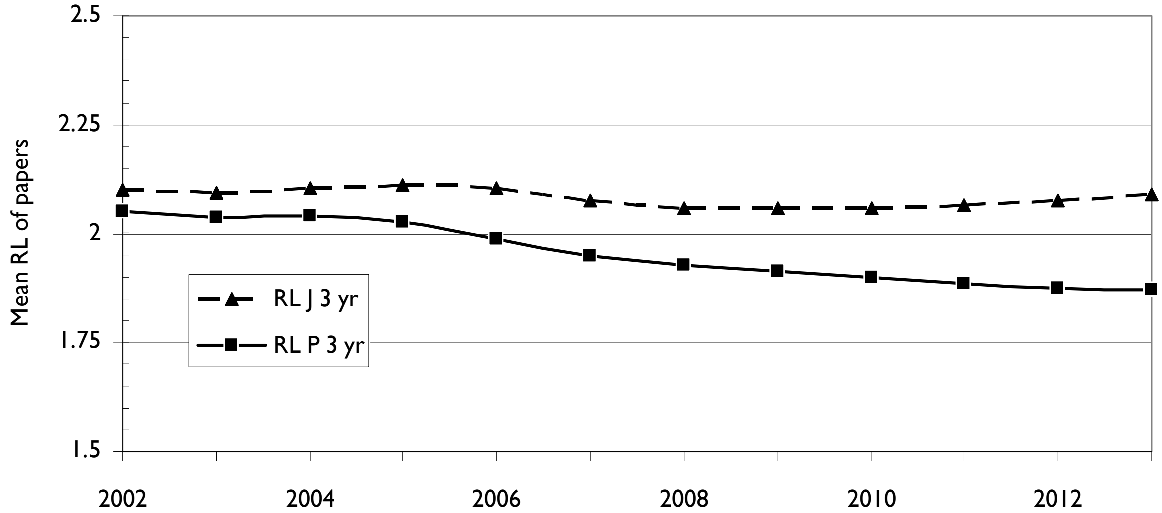


Figure 4. Mean research level of cancer papers from EUR31 countries, 2002-13. RL = 1 is clinical; RL = 4 is basic research

Table 5. Citation performance of EUR31 countries in 2002-09, ranked by the percentage with 53 or more cites in the five years following publication (ACI) (Top) rather than the mean value.

ISO	*Mean*	*Top*	*%*		*ISO*	*Mean*	*Top*	*%*		*ISO*	*Mean*	*Top*	*%*
CH	19.1	280.1	6.67		FR	14.1	763.0	4.12		CZ	9.5	27.4	1.66
NL	19.4	603.1	6.17		ES	14.2	366.3	4.11		BG	6.3	3.3	1.27
UK	18.0	1469.1	6.14		IT	14.3	905.5	3.96		PL	7.9	50.9	1.25
IS	19.3	6.9	5.83		IE	13.8	47.0	3.74		RO	6.0	4.3	1.05
BE	17.2	216.6	5.44		NO	15.0	86.3	3.61		LT	5.8	1.2	1.05
DK	17.5	139.2	5.30		LV	9.4	1.5	3.25		SI	7.3	4.1	0.83
FI	16.6	117.4	4.74		PT	12.6	30.9	3.17		EE	8.5	0.3	0.60
SE	15.6	267.7	4.51		GR	9.5	89.9	1.93		MT	3.9	0.1	0.50
AT	15.0	158.0	4.37		CY	9.3	0.7	1.89		HR	5.1	3.7	0.47
LU	16.7	2.4	4.26		HU	9.3	21.7	1.81					
DE	14.3	1211.5	4.22		SK	8.9	7.7	1.75					

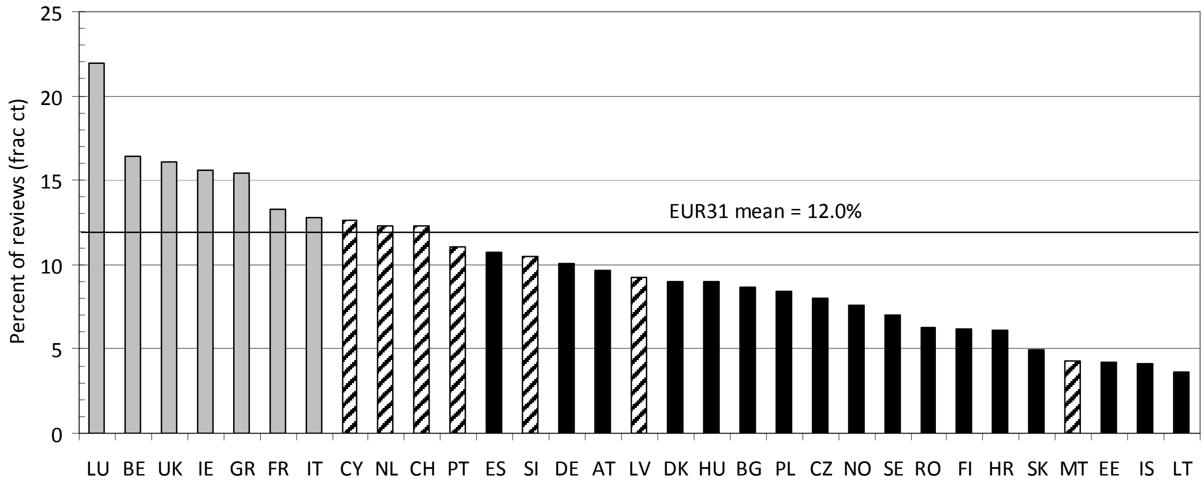


Figure 5. Percentage of reviews among cancer research papers of EUR31 countries, 2002-13.

Bars shaded light grey: significantly above EUR31 average; bars in black: significantly below EUR31 average; striped bars: not significantly different from average.

Types of research

The numbers and percentages of papers of the 11 research types listed in Table 2 are shown in Table 6 for the 10 leading European countries in terms of fractional count output.

Table 6. Outputs of papers from 10 leading European countries in 11 different types of cancer research, 2002-13, and their relative commitment compared with the European average. Countries ordered by their fractional count totals.

Values > 2 shown in bold and large type, values > 1.41 in bold, values < 0.71 in *italics*, values < 0.5 in *small italics*.

	GENE	*CHEM*	*PROG*	*SURG*	*PATH*	*EPID*	*RADI*	*DIAG*	*SCRE*	*PALL*	*QUAL*
EUR	48259	28240	27189	26585	19119	12836	12085	11334	3437	3152	1369
%	19.1	11.2	10.8	10.5	7.6	5.1	4.8	4.5	1.4	1.2	0.5
DE	1.03	0.88	0.94	1.09	1.09	*0.69*	1.07	1.12	*0.61*	*0.70*	0.82
IT	0.86	1.35	0.94	1.17	0.96	0.89	0.73	0.93	0.74	0.84	*0.49*
UK	0.96	0.84	1.02	1.03	0.91	1.10	0.97	0.92	1.32	**1.62**	**1.50**
FR	0.85	1.11	0.92	1.06	0.88	0.94	1.17	0.92	0.89	*0.65*	*0.58*
NL	0.97	0.99	1.13	1.07	0.99	1.28	1.76	1.03	**2.07**	1.32	**2.24**
ES	1.12	1.08	1.05	0.75	1.07	0.92	*0.60*	1.11	0.93	0.89	0.78
SE	1.31	0.74	1.31	0.67	0.78	**2.26**	1.08	0.90	1.17	**1.79**	**1.92**
PL	1.28	0.91	0.73	0.74	0.98	0.94	0.83	1.00	*0.45*	0.74	*0.64*
GR	0.99	**1.42**	1.01	1.08	0.98	0.79	*0.70*	0.97	*0.65*	0.94	0.87
CH	0.86	0.94	0.91	0.92	1.08	*0.64*	1.22	1.27	*0.61*	0.74	*0.62*

Screening, palliative care and quality of life research receive little attention, although the UK is prominent in the latter two, as are Sweden and the Netherlands.

Research on different cancer sites

The table below shows values only for the 13 leading cancer sites, but they account for 86% of the papers on any one of the 22 sites listed in Table 3. Breast cancer is the site of greatest research interest, but it accounts for fewer DALYs than colorectal cancer and many fewer than lung cancer in all European countries.

Table 7. Outputs of papers from 10 leading European countries on 13 different cancer sites, 2002-13, and their relative commitment compared with the European average. Countries ordered by their fractional count totals.

Values > 1.41 shown in bold, values < 0.71 in *italics*, values < 0.5 in *small italics*.

	MAM	*COL*	*LEU*	*LYM*	*PRO*	*LUN*	*LIV*	*STO*	*BRA*	*MEL*	*MOU*	*KID*	*OVA*
%	9.15	6.17	5.26	4.22	4.04	3.68	3.46	3.40	3.39	3.28	2.33	1.95	1.93
DE	0.75	0.85	1.07	1.05	1.06	0.81	1.16	1.19	1.24	1.12	1.04	1.22	0.78
IT	0.92	0.93	1.08	1.09	0.89	1.15	1.34	1.03	1.10	1.04	0.82	0.90	1.06
UK	1.19	1.15	0.92	0.87	1.09	0.80	*0.68*	0.77	0.83	0.85	1.23	0.83	1.08
FR	0.92	0.89	1.00	1.14	0.99	1.14	1.24	0.92	0.99	0.89	*0.61*	1.35	0.85
NL	1.09	1.33	0.81	0.75	1.10	1.20	0.80	1.03	0.80	1.00	1.52	0.86	0.84
ES	0.99	1.10	0.99	1.25	0.78	1.24	1.18	0.97	1.00	0.95	1.19	1.05	*0.69*
SE	1.17	1.14	1.07	0.87	**1.61**	*0.67*	*0.52*	0.83	1.12	0.85	*0.71*	0.74	0.96
PL	1.01	0.93	**1.43**	*0.71*	*0.47*	1.19	*0.62*	1.16	0.83	1.07	*0.64*	0.96	**1.93**
GR	1.23	1.07	0.88	1.25	0.82	**1.56**	1.03	1.40	*0.71*	*0.67*	1.15	0.85	**1.64**
CH	0.84	0.72	0.73	1.20	0.90	1.01	0.94	*0.65*	1.22	1.38	1.28	0.81	0.63

The correlation between country relative research commitment, as shown above, and its disease burden in DALYs is poor for almost all countries, and is only positive for Italy and France. Prostate cancer, which accounts for almost as many deaths among men as breast cancer does among women, receives far less research attention, although Sweden has a rather high relative commitment to research on the site.

Conclusions

The tables and figures shown here provide a sample of the data available in the full database, whose details will be made available in reports to the European Commission. But already it is clear that the European cancer research portfolio, although very extensive, does not reflect the burden of disease in the different countries, and that some research types are relatively neglected by most countries – in particular, those pertaining to end-of-life issues such as palliative care and quality of life. There may well be interactions with mental disorders, and there is an increasing amount of research that covers both these two sub-fields (Purushotham *et al*., 2013) but the clinical need underlying this research has not been sufficiently explored. There is also a relative paucity of research on radiotherapy and surgery, compared with chemotherapy, and of research on lung cancer, which causes the greatest burden of disease from cancer in Europe.

Acknowledgments

This project is being funded by the European Commission under the Sixth Framework Programme. We are very grateful for the macros written by Dr Philip Roe of Evaluametrics Ltd, that enabled the analysis to be carried out swiftly and accurately.

References

Bailey, LJ Sanson-Fisher, R Aranda, S D'Este, C Sharkey & K Schofield, P (2010) Quality of life research: types of publication output over time for cancer patients, a systematic review *European Journal of Cancer Care,* 19 (5) , 581-588.

Brolmann, FE Groenewold, MD Spijker, R vanderHage, JA Ubbink, DT & Vermeulen, H (2012) Does evidence permeate all surgical areas equally? publication trends in wound care compared to breast cancer care: a longitudinal trend analysis *World Journal of Surgery* 36 (9), 2021-2027.

Campbell, D Picard-Aitken, M Cote, G Caruso, J Valentim, R Edmonds, S *et al.* (2010) Bibliometrics as a performance measurement tool for research evaluation: the case of research funded by the National Cancer Institute of Canada *American Journal of Evaluation*, 2010, 31 (1), 66-83.

Glynn, RW Scutaru, C Kerin, MJ & Sweeney, KJ (2010) Breast cancer research output, 1945-2008: a bibliometric and density-equalizing analysis *Breast Cancer Research*, 12 (6).

Glynn, RW Lowery, AJ Scutaru, C O'Dwyer & T Keogh, I (2012) Laryngeal cancer: quantitative and qualitative assessment of research output, 1945-2010 *Laryngoscope* 122(9), 1967-1973.

Grossi, F Belvedere, O & Rosso, R (2003) Geography of clinical cancer research publications from 1995 to 1999 *European Journal of Cancer* 39(1), 106-111.

Healy, NA Glynn, RW Scutaru, C Groneberg, D Kerin, MJ & Sweeney, KJ (2011) The h index and the identification of global benchmarks for breast cancer research output *Breast Cancer Research and Treatment* 127(3), 845- 851.

Ho, YS Nakazawa, K Sato, S Tamura, T Kurishima, K & Satoh, H (2013) Cisplatin for small cell lung cancer: associated publications in Science Citation Index Expanded *Oncology Letters* 5(2), 684-688.

Kunath, F Krause, SF Wullich, B Goebell, PJ Engehausen, DG Burger, M *et al*., (2013) Bladder cancer , the neglected tumor: a descriptive analysis of publications referenced in MEDLINE and data from the register clinicaltrials.gov *BMC Urology* 13.

Lewison, G (2003) The publication of cancer research papers in high impact journals, *ASLIB Proceedings* 55(5-6), 379-387.

Lewison, G & Paraje, G (2004) The classification of biomedical journals by research level. *Scientometrics,*60(2) , 145-157.

Murray, CJL, Ezzati, M, Flaxman, AD, Lim, S, Lozano, R, Michaud, C *et.al*. (2012) GBD 2010: design, definitions and metrics *The Lancet*, 380 (9859) 2063-2066.

Perez-Santos, JLM & Anaya-Ruiz, M (2013) Mexican breast cancer research output, 2003-2012 *Asian Pacific Journal of Cancer Prevention* 14(10), 5921-5923.

Purushotham, A Bains, S Lewison, G Szmukler, G & Sullivan, R (2013) Cancer and mental health – a clinical and research unmet need *Annals of Oncology*, 24, 2274-2278.

Skovlund, E (1998) A critical review of papers from clinical cancer research *Acta Oncologica*, 37(4), 339-345.

Ugolini, D Parodi, S & Santi, L (1997) Analysis of publication quality in a cancer research institute *Scientometrics* 38(2), 265-274.

Zhang, XC Huang, DS & Li, F (2011) Cancer nursing research output and topics in the first decade of the 21st century: results of a bibliometric and co-word cluster analysis *Asian Pacific Journal of Cancer Prevention* 12(8), 2055-2058.

Library and Information Science Research Output: A study based on Web of Science

Nabi Hasan* and Mukhtiar Singh**

*Deputy Librarian, Indian Institute of Technology Delhi - 110016, India, Phone- +91-9560978667
hasan@library.iitd.ac.in

**Principal Technical Officer (Library), CSIR-Institute of Himalayan Bioresource Technology, Palampur (HP) - 176061, India
msingh@ihbt.res.in

Abstract

The paper attempts to evaluate the trend of world literature on "Information and Library Science" (LIS) in terms of the output of research publications as indexed in the Science Citation Index, Social Sciences Citation Index (SSCI) and Arts & Humanities Citation Index (A&HCI) during the period from 1975 to 2012. An overall total of 311,886 records were retrieved on Library and Information Science including all forms of literature. The records were categorized under 23 types of documents. A scientometric assessment of the status of research papers is presented in the study by way of analyzing some of the features of publications of the study period; Year-wise distribution of publications on Library and Information Science, Form-wise distribution, Language-wise distribution, Annual output of publications, Geographical distribution, Subject dispersion, Institutional distribution, Sources preferred for publishing, Indian contribution to LIS, etc. The average number of papers published per year was 2291.53. The highest number of papers, i.e. 3.94% were published in the year 2011. United States contributed highest to the tune of 48.47% and India was at eleventh position with 0.95% research papers. Most productive institution was University of California, which contributed a total of 2.06% research papers. The 'Scientist' journal was at the first position by publishing 4.63% research papers. English language with 91.20% research papers was at the top. The paper has given special emphasis for Indian LIS research output. The study may be useful to subject specialists, analysts, researchers, students and policy makers to look into the trends and make effective policies on the basis of inferences drawn in this paper.

Keywords: Information and Library Science; Web of Science; Research output; Scholarly output, Scientometric analysis; Mapping research-Library Science; Library and Information Science Research - India

Introduction

Library and Information Science is an important applied discipline in which lot of literature is being produced. It is very important to map the dynamics of the subject. Web of Science is an important database which cover important journals of the discipline. Such a study may be very useful to map research contributions in the discipline and will enable to understand the strengths and weaknesses of the subject. The study also covers the contributions from Indian authors in LIS and inferences drawn on the basis of this study may be very useful.

Objectives

The main objective of this study is to quantify the trend of world literature on "Library and Information Sciences" and to make the scientometric assessment of the status of research by way of analyzing the following features of publications of the study period:

i) Year-wise distribution of publications on Library and Information Science
ii) Form-wise distribution
iii) Language-wise distribution
iv) Annual output of publications
v) Geographical distribution
vi) Subject dispersion
vii) Institutional distribution

viii) Sources preferred for publishing
ix) Indian contribution to LIS, etc.

Materials, Methods and Limitations

Data with the keyword "library and information science" was collected from the online Web of Science-Science Citation Index, Social Sciences Citation Index (SSCI) and Arts & Humanities Citation Index (A&HCI) of the period from 1975 to 2012. This database was chosen as it is most reputed and very comprehensive database covering all aspects of subject under study. It facilitates quick, powerful access to the bibliographic and citation information on world scientific literature. Publications were searched with Web of Science category of 'Information & Library Science' in basic index by limiting the period to 1975-2012. A total of 311,886 records were retrieved in LIS covering all forms of literature, and further filtered/restricted with in these records to callout 87,078 research papers selecting Articles and Reviews only for analysis and interpretation.

Data retained was scanned to facilitate Year-wise distribution of publications on Library and Information Science, Form-wise distribution, Language-wise distribution, Annual output of publications, Geographical distribution, Subject dispersion, Institutional distribution, Sources preferred for publishing, Indian contribution to LIS, etc. Further analyses was done by using the Microsoft's Excel and inferences were drawn.

The proposed study has its own limitations as every field has its limits. The study was confined to the specific subject- Information & Library Science. Publications of research value, published and indexed during 1975 to 2012 were considered for the data collection and analysis. The Web of science was preferred because it is considered a quality database in selection of resources for coverage of the subject. It indexes more than 6,650 major journals across 150 scientific disciplines and includes all cited references captured from indexed articles (http://apps.webofknowledge.com).

Results and Discussion

A total of 311, 886 records as shown year-wise in **Table 1**, were retrieved on Library and Information Science. The records have been categorized under 23 types of documents as depicted in **Table 2**. Maximum number of records were indexed as Book Reviews 145,772 (46.74%) followed by Articles with 84,773 (27.18%) and Proceeding Papers with 36,869 (11.82%). Moreover, these records were communicated in 20 languages of the world as mentioned in **Table 3**. English language counted highest number of records with 298,296 (95.64%) followed by German with 6,686 (2.14%) records and Russian was at third position with 3563 (1.14%) records. **Being the primary source of information, Research articles and Reviews 87,078 (27.92%) only were taken for further analysis**. The data was downloaded and analyzed keeping in view the objectives of the study.

Table 1. Year-wise distribution of publications on LIS during 1975 to 2012 (all forms/total records)

Column- I			*Column- II*		
Year	*No. records*	*% age*	*Year*	*No. records*	*%age*
2012	9067	2.91	1993	8876	2.85
2011	10370	3.33	1992	6455	2.07
2010	10574	3.39	1991	6970	2.24
2009	10359	3.32	1990	7855	2.52
2008	13686	4.39	1989	6134	1.97
2007	13311	4.27	1988	5315	1.70

2006	10808	3.47	1987	5510	1.77
2005	10830	3.47	1986	4495	1.44
2004	10561	3.39	1985	4291	1.38
2003	11113	3.56	1984	4154	1.33
2002	10857	3.48	1983	4522	1.45
2001	11276	3.62	1982	4043	1.30
2000	11769	3.77	1981	3840	1.23
1999	12057	3.87	1980	3952	1.27
1998	12417	3.98	1979	3923	1.26
1997	13255	4.25	1978	3819	1.22
1996	11687	3.75	1977	3514	1.13
1995	12828	4.11	1976	2642	0.85
1994	12269	3.93	1975	2482	0.80

Table 2. Form-wise distribution of records on LIS (all forms/total records)

Type of document	*No. of records*	*% age*
Book Review	145772	46.74
Article	84773	27.18
Proceedings Paper	36869	11.82
Editorial Material	21707	6.96
Letter	10175	3.26
News Item	3950	1.27
Meeting Abstract	3116	1.00
Review	2305	0.74
Software Review	2281	0.73
Note	2064	0.66
Biographical Item	745	0.24
Item About An Individual	666	0.21
Database Review	616	0.20
Correction Addition	548	0.18
Bibliography	545	0.18
Correction	487	0.16
Reprint	204	0.07
Discussion	188	0.06
Hardware Review	90	0.03
Chronology	7	0.00
Poetry	3	0.00
Fiction Creative Prose	2	0.00
Film Review	1	0.00

Table 3. Language-wise distribution of records on LIS (all forms/total records)

Year	*No. records*	*% age*
English	298296	95.64
German	6686	2.14
Russian	3563	1.14
French	1187	0.38
Spanish	968	0.31
Portuguese	587	0.19
Japanese	348	0.11
Hungarian	145	0.05
Chinese	81	0.03
Italian	7	0.00
Swedish	4	0.00
Estonian	3	0.00
Multiple languages	3	0.00
Polish	3	0.00
Welsh	3	0.00
Romanian	2	0.00
Catalan	1	0.00
Czech	1	0.00
Georgian	1	0.00
Rumanian	1	0.00

Annual output of publications on LIS

During 1975-2012, a total of 87,078 research papers were published on LIS by various countries in the world. The average number of publications produced per year was 2291.53 papers. The highest number of publications 3427 (3.94%) were produced in 2011. **Table 4** gives annual growth rate in LIS. It can be clearly visualized from the figure that growth of the literature was very slow with ups and down. Papers contributed during 1975 were 1369 (1.57%) which later increased to 2000 by 68.45% growth in 1985. There was a slight increase of 17.08% accounting 2412 research papers in 1995 in a period of ten years. In next ten year (1995-2005), research papers increased by 8.42% with 2615 papers which is even less then the earlier decade's growth. Highest growth was recorded in 2011 with 3427 research papers which is 4.17% of the total papers indexed.

Table 4. Year-wise distribution of research papers on LIS (Articles and Reviews only)

Column-I			*Column-2*		
Year	*No. of publications*	*% age*	*Year*	*No. of publications*	*% age*
2012	3364	3.86	1993	2083	2.39
2011	3427	3.94	1992	2105	2.42
2010	3283	3.77	1991	2201	2.53
2009	3182	3.65	1990	2234	2.57
2008	2992	3.44	1989	1966	2.26
2007	2859	3.28	1988	1991	2.29
2006	2609	3.00	1987	2087	2.40
2005	2615	3.00	1986	1956	2.25

2004	2193	2.52	**1985**	2000	2.30
2003	2331	2.68	1984	1979	2.27
2002	2431	2.79	1983	1938	2.23
2001	2394	2.75	1982	1943	2.23
2000	2316	2.66	1981	1902	2.18
1999	2352	2.70	1980	1929	2.22
1998	2317	2.66	1979	1987	2.28
1997	2322	2.67	1978	2013	2.31
1996	2329	2.68	1977	1686	1.94
1995	2412	2.77	1976	1514	1.74
1994	2467	2.83	**1975**	1369	1.57

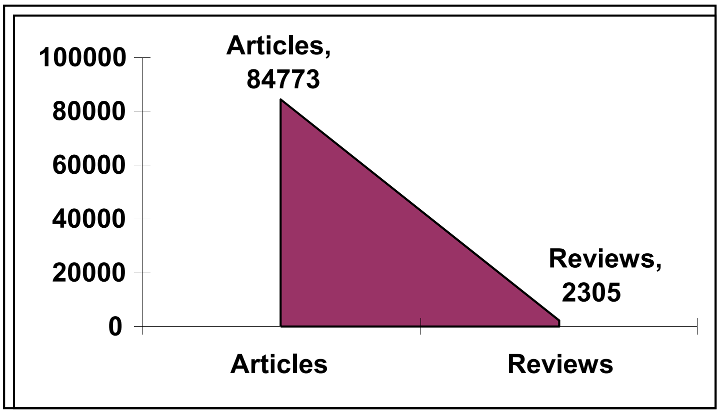


Figure 1. Form of papers contributed in LIS during 1975-2012 (Articles/Reviews only)

Language-wise distribution

Language of communication used for publication of research papers plays a major role in highlighting and publicizing the research output. English was found to be on top of the list with 79414 (91.20%) publications followed by Russian with 2796 (3.21%), German 2391 (2.75%), French 809 (0.93 %), Spanish 779 (0.90%), Portuguese 475 (0.55), Japanese 315 (0.36) and Hungarian 92 (0.11%). Other languages were meager in contributions as shown in **Table 5**. Publications reported were in 12 languages during the study period.

Table 5. Language-wise distribution of papers

Language of publications	*No. of publications*	*% age*
English	79414	91.20
Russian	2796	3.21
German	2391	2.75
French	809	0.93
Spanish	779	0.90
Portuguese	475	0.55
Japanese	315	0.36
Hungarian	92	0.11
Chinese	4	0.01
Georgian	1	0.00
Romanian	1	0.00
Welsh	1	0.00

Geographical distribution

Growth of publications in LIS was illustrated in table 4 of the period from 1975-2012. The cumulative publication productivity of top 20 out of 185 countries involved in LIS during 1975-2012 is given in **Table 6.** Amongst these countries, the USA topped the list with 42210 (48.47%) publications, followed by UK with 7124 (8.18%), Canada 3835 (4.40%), Germany 1920 (2.21%), Australia 1710 (1.96%), Spain 1658 (1.90%), France 1584 (1.82%), China 1447 (1.66%), the Netherlands 1470 (1.69%), Taiwan 857 (0.98%) and India with 830 (0.95%) ranked at 11th place in terms of research publication output during 1975-2012. Contribution in library and information science was from 185 countries and top 20 countries contributed 71250 (81.08%) research papers and 165 countries contributed only 15,828 (18.20%) research papers.

Table 6. Top twenty contributing countries in LIS

Name of country	*No. of research papers*	*% age*
USA	42210	48.47
UK	7124	8.18
Canada	3835	4.40
Germany	1920	2.21
Australia	1710	1.96
Spain	1658	1.90
France	1584	1.82
Netherlands	1470	1.69
China	1447	1.66
Taiwan	857	0.98
India	830	0.95
Japan	797	0.92
Germany	775	0.89
Scotland	769	0.88
Belgium	754	0.87
USSR	744	0.85
Brazil	729	0.84
South Korea	726	0.83
Italy	717	0.82
Singapore	594	0.68

Subject Dispersion

The narrower subjects have been further addressed in research publications under 21 broad subject categories during 1975-2012. It was observed that maximum number of publications were in Computer Science with 31753 (36.47%) which indicates the application of ICTs in the field of library management and services, followed by Business economics 4650 (5.34%), Science technology and other topics 4030 (4.63%), communication 3197 (3.67%), Medical Informatics 2765 (3.18%), Telecommunications 1476 (1.76%) and papers related to other remaining categories have been shown in the **Table 7**.

Table 7. Top ten sub-subject categories in research papers in addition to LIS

Subject category	*No. of publications*	*% age*
Computer science information systems	28044	32.21
Computer science interdisciplinary applications	6386	7.33
Management	4650	5.34
Multidisciplinary sciences	4030	4.63
Communication	3197	3.67
Medical informatics	2765	3.18
Social sciences interdisciplinary	1919	2.20
Telecommunications	1476	1.70
Law	1219	1.40
Geography	1132	1.30

Institutional Distribution

Table 8 shows ranking of top institutions which contributed 500 or more papers on LIS. The institutions from USA performed excellent and secured all top positions. University of California System USA topped with 1797 (2.06%) papers followed by University of Illinois System, USA with 1373 (1.58%) and State University of New York, USA is at third position with 1033 (1.19%).

Table 8. Highly productive Institutions in LIS

Name of organization	*No. of records*	*% age*
University of California System, USA	1797	2.06
University of Illinois System, USA	1373	1.58
State University of New York, USA	1033	1.19
Indiana University, USA	991	1.14
University of Illinois Urbana Champaign , USA	955	1.10
University of Wisconsin System , USA	798	0.92
University of London, UK	706	0.81
University of North Carolina Chapel Hill, USA	667	0.77
Harvard University, USA	651	0.75
University of Michigan, USA	635	0.73
University of Pittsburgh, USA	608	0.70
University of California Los Angeles, USA	606	0.70
Penn State University, USA	605	0.70
Rutgers State University, USA	587	0.67
Loughborough University, USA	543	0.62
University of Washington, USA	539	0.62
University of Washington Seattle, USA	537	0.62
Columbia University, USA	510	0.59

Sources preferred for Publishing

Table 9 provides the list of top 20 productive journals preferred for publication in LIS. The journal Scientist was at the top with 4030 (4.63%) papers followed by Library Journal with 3805 (4.37%) and Scientometrics was at third position with 3012 (3.46%). These top journals contributed 43.00% of the total papers during 1975-2012. Of these 20 journals, the first 10 journals contributed 22027 papers (26.99%), next 10 journals contributed 13,072 papers (16.01%), and rest 46523 (57%) papers were published in 151 journals in the cumulative output. The average papers published by all the 185 journals were 477.32 papers per journal in the total of 81622 papers.

Table 9. Top journals in the field of LIS preferred for publications

Name of journal	*No. of papers*	*% age*
Scientist	4030	4.63
Library Journal	3805	4.37
Scientometrics	3012	3.46
Journal of the American Medical Informatics Association	2643	3.04
Information Processing Management	1943	2.23
Journal of the American Society for Information Science and Technology	1821	2.09
Online	1740	2.00
Library Trends	1583	1.82
Aslib Proceedings	1571	1.80
Journal of Academic Librarianship	1541	1.77
Information Management	1523	1.75
Journal of the American Society for Information Science	1516	1.74
Nauchno Tekhnicheskaya Informatsiya Seriya 1 Organizatsiya I Metodika Informatsionnoi Raboty	1501	1.72
Telecommunications Policy	1476	1.70
Journal of Information Science	1353	1.55
College Research Libraries	1335	1.53
Nauchno Tekhnicheskaya Informatsiya Seriya 2 Informatsionnye Protsessy I Sistemy	1295	1.49
Data Management	1259	1.45
Social Science Information Sur Les Sciences Sociales	1197	1.38
Wilson Library Bulletin	1152	1.32

India: Contribution to LIS

During 1975–2012, a total of 830 papers were published by Indians in the area of LIS which is only 0.95% contribution to the global literature on the subject. The average number of publications produced per year are 21.84 papers. The highest number of publications 50 (6.20%) were produced in 2011. **Table 10** gives annual growth rate of papers contributed in LIS from India. It can be clearly visualized from the figure that growth of the literature was more or less similar to the growth of world. Papers contributed during 1975 were 5 (0.60%). Out of total 830 papers, 20 (1.20%) were reviews.

Table 10. Year-wise contribution of articles and reviews from India

Column-I			*Column-II*		
Year	*No. of papers*	*% age*	*Year*	*No. of papers*	*% age*
2012	41	4.94	1993	16	1.93
2011	50	6.02	1992	15	1.81
2010	48	5.78	1991	14	1.69
2009	40	4.82	1990	9	1.08
2008	38	4.58	1989	21	2.53
2007	23	2.77	1988	18	2.17
2006	33	3.98	1987	18	2.17
2005	30	3.61	1986	14	1.69
2004	20	2.41	1985	19	2.29
2003	23	2.77	1984	14	1.69
2002	24	2.89	1983	12	1.45
2001	22	2.65	1982	7	0.84
2000	17	2.05	1981	30	3.61
1999	20	2.41	1980	19	2.29
1998	27	3.25	1979	19	2.29
1997	12	1.45	1978	38	4.58
1996	11	1.33	1977	15	1.81
1995	19	2.29	1976	17	2.05
1994	12	1.45	1975	5	0.60

India: Collaborations in Papers

Table 11 provides an overview of collaboration of Indian authors with the authors of other countries in LIS. Indian authors collaborated for 140 (16.87%) papers. The highest collaboration was between Indian and USA authors with 59 (7.11%) papers followed by Belgium with 14 (1.69%) and UK with 9 (1.08%) papers. There were 32 countries collaborated with Indian authors.

Table 11. Status of foreign collaboration in contributions on LIS from India

Name of country	*No. of papers*	*% age*
USA	59	7.11
Belgium	14	1.69
UK	9	1.08
Germany	7	0.84
Singapore	6	0.72
Netherlands	4	0.48
China	4	0.48
Bangladesh	3	0.36
Finland	3	0.36
France	3	0.36
Australia	2	0.24
Brazil	2	0.24
Iran	2	0.24
Kenya	2	0.24

Switzerland	2	0.24
Thailand	2	0.24
Antigua Barbu	1	0.12
Bulgaria	1	0.12
Canada	1	0.12
Denmark	1	0.12
Egypt	1	0.12
Ethiopia	1	0.12
Hong Kong	1	0.12
Malaysia	1	0.12
Nigeria	1	0.12
Norway	1	0.12
Qatar	1	0.12
South Africa	1	0.12
South Korea	1	0.12
Swaziland	1	0.12
Taiwan	1	0.12
Yugoslavia	1	0.12

India: Top Contributing Institutions

Table 12 shows position of top ten institutions from India as per the contribution of papers in LIS. CSIR topped the position with 149 (17.96%) papers followed by Indian Statistical Institute with 53 (6.39%) papers, Indian Institute Technologies with 41 (4.94%) and Indian Institute of Management with 31 (3.74%) papers. Contributions were from a total of 460 organizations.

Table 12. Top contributing institutions in LIS from India

Name of Institution	*No. of papers*	*% age*
CSIR	149	17.96
Indian Statistical Institute	53	6.39
Indian Institute Technologies	41	4.94
Indian Institute of Management	31	3.74
Indian Institute of Science	20	2.41
University of Delhi	20	2.41
Karnataka University	19	2.29
Bhabha Atomic Research Centre	17	2.05
Guru Nanak Dev University	14	1.69
University of Mysore	13	1.57

India: Prolific Authors

Table 13 shows most productive authors in LIS from India, who contributed ten or more papers. Highly productive author was Gupta, BM with contribution of 31 (3.74%) papers followed by Garg, KC with 30 (3.61%) papers and Rao, IKR at third position by contributing 21 (2.53%) papers. A total of 865 authors contributed papers and an average of 1.04 papers contributed by a single author.

Table 13. Prolific authors from India

Name of author	*No. of papers*	*% age*
Gupta, BM	31	3.74
Garg, KC	30	3.61
Rao, IKR	21	2.53
Arunachalam, S	20	2.41
Kumar, S	20	2.41
Neelameghan, A	19	2.29
Prathap, G	14	1.69
Rao, SS	14	1.69
Satija, MP	14	1.69
Kumar, V	13	1.57
Nagpaul, PS	12	1.45
Sengupta, IN	12	1.45
Bhattacharya, S	11	1.33
Basu, A	10	1.21
Karisiddappa, CR	10	1.21

India: Preference of Journals for Publications

Table 14 provides distribution of articles in journals having published 10 or more papers during 1975-2012. A total of 79 journals published papers from India. Out of total 79 journals, the leading journals preferred by the LIS authors were; Scientometrics 180 (21.69%) followed by Library Science with a Slant to Documentation with 92 (11.08%) and Electronic Library 59 (7.11%). On the whole an average number of 10.51 papers published by a single journal.

Table 14. Preferred journals for publication of papers by Indian authors

Name of journal	*No. of papers*	*% age*
Scientometrics	180	21.69
Library Science with a Slant to Documentation	92	11.08
Electronic Library	59	7.11
International Library Review	44	5.30
Program: Electronic Library and Information Systems	40	4.82
LIBRI	25	3.01
Information Processing Management	22	2.65
International Forum on Information and Documentation	20	2.41
Journal of Information Science	18	2.17
International Classification	17	2.05
International Journal of Information Management	15	1.81
Knowledge Organization	15	1.81
Malaysian Journal of Library and Information Science	14	1.69
Journal of Knowledge Management	13	1.57
Online Information Review	11	1.33
International Journal of Computer Information Sciences	10	1.21
Social Science Information Sur Les Sciences Sociales	10	1.21
Telecommunications Policy	10	1.21

Conclusion

During the 38 year period under study, growth of research papers was found to be inconsistent and have many ups and downs. However the overall trend of the period is highly fluctuating but in increasing order. Evaluation of research performance in terms of research publications is considered as an integral part of science, and important in the scientific community in the field of LIS. The present paper has carried out an evaluation of publications on LIS of about four and the important findings though already elaborated in the analysis part above of 38 years; some of them have been highlighted below as conclusion:

i) USA is the major contributor with 42210 (48.47%) papers to its credit on the subject. Growth of the literature on LIS showed increasing trend during 1975-2012. Only twenty countries have contributed 81.08% of research papers.
ii) Most productive institution was University of California by contributing 1797 (2.06%) publications followed by the University of Illinois System with (1.58%) research papers.
iii) Of the two types of literature covered in the study, 84773 (87.65%) of the literature contain Articles followed by Reviews with 2305 (2.65%).
iv) The most preferred source was the journal - Scientist with 4.63% followed by Library journal with 4.37% publications.
v) English language dominated with 91.20% publications and other languages like, Russian, German, French and Spanish did also find place in publications on information and library science.
vi) In sub-subject categorization of publications, computer science and information systems was at first position with 32.21% publications.
vii) In respect to the India, a total of 830 papers were contributed in the area of LIS during the period which is 0.95% contribution to the global published literature. The average number of papers produced per year was 21.84 papers. CSIR topped the position with 149 (17.96%) papers followed by Indian Statistical Institute with 53 (6.39%) papers. Highly productive author from India was Gupta, BM with contribution of 31 (3.74%) papers followed by Garg, K.C. with 30 (3.61%) papers and Rao, IKR at third position by contributing 21 (2.53%) papers.
viii) The study provides a base for researchers, institutions and policy makers to initiate new research projects or studies in the area. It also makes researchers aware regarding the growth of quality research literature of their interest in the subject.
ix) It helps in identifying highly productive countries and researchers in the area. The study provides details of institutions and researchers working in specific subject areas in the field of library science.
x) The study may help policy makers to make effective policies related to library and information science.

References

Web of Science database (2013). Retrieved June 3, 2014 from: http://apps.webofknowledge.com

Gupta, B.M. & Kaur, Har (2013). World glaucoma research: a quantitative analysis of research output during 2002-11. *Annals of Library and Information Studies*, 60, 98-106.

Mahesh, G. & Wadhwa, N.K. (2013). Web of Science based ranking of Indian library and information science journals. *Collnet Journal of Scientometrics and Information Management*, Online first, available at: http://www.tarupublications.com/journals/cjsim/FullText/new%20pdf/05_CJSIM6-2.pdf (accessed 3 June 2014).

Hasan, Nabi & Sewa Singh (2007). Agricultural Research in Himachal Pradesh: A Profile based on AGRICOLA, AGRIS, CAB and FSTA CD-ROM databases. *SRELS Journal of Information Management,* 44, 279-300.

Robertson, S., Jitan, S. & Reese, K. (1997). Web-based collaborative library research. *Proceedings of the 2nd ACM International Conference on Digital Libraries,*152-160.

Singh, M. and Hasan, Nabi (2013). Intellectual Property: a scientometric study of scholarly output. *Journal of Knowledge and Communication Management,* 3, 13-28.

Hasan, Nabi and Singh, Mukhtiar (2013). Digital Library Literature: A Scientometric analysis. In Ganguly, Shantanu, et. al. (Eds), *International Conference on Digital Libraries-2013 on "Vision 2020: Looking back 10 years and forging new frontiers (*ICDL-2013): Proceedings (pp, 319-329) New Delhi: TERI Press.

The Mutual Role of Scientometrics and Foresight

Roya Pournaghi* and Leila Nemati-Anaraki**

*Assistant Professor of Iranian Research Institute for Information Science & Technology, Tehran, Iran
pournaghi@irandoc.ac.ir

**Iran University of Medical Sciences, Tehran, Iran
lnemati@yahoo.com

Abstract

Foresight is a science and art which helps human being to know the future events, opportunities, and threats and to wisely choose the optimal (possible) futures from among possible (exploratory) ones and to consider future as an uncertain and instable setting. Foresight is the process of developing a range of views of possible ways in which the future could develop, and understand these sufficiently well to be able to decide what decisions can be taken today to create the best possible tomorrow. Scientometrics addresses science measurement in different areas and its overall objective is measuring the creation, distribution and consumption of science both quantitatively and qualitatively. Both scientometrics and foresight are interdisciplinary and are quite similar in observing scientific activities, analyzing the trends, scientific prospective and predictions. Indeed, making use of foresight means taking action for making policies in order to realize an optimal future through which a number of other helpful results would be obtained: developing new networks and communications, goal-setting and enhancing the sense of shared commitment, publishing the information, and prioritizing the options and actions. In this article, the authors try to examine and elaborate the mutual relationship between scientometrics and futurology in different scientific texts then they will discuss the status and role of scientometrics as a tool for foresight-related studies.

Keywords: Scientometrics, foresight, Futurology, Information Sciences

Introduction

The contemporary world is the setting of the growing, drastic and dynamic changes. The changes are so unpredictable and unexpected that even a minute amount of ignorance could lead to the huge strategic shock in all political, social, and even cultural areas. In such a changing, instable and uncertain setting, the only approach which might have more chance of success is attempting to build future. Although making an attempt in this regard is always juxtaposed with high risk-taking, undertaking it seems to be more logical than merely observing the future changes.

Foresight is a science and art which helps human being to know the future events, opportunities, and threats and to wisely choose the optimal (possible) futures from among possible (exploratory) ones and to consider future as an uncertain and instable setting. Foresight is the process of developing a range of views of possible ways in which the future could develop, and understanding these sufficiently well to be able to decide what decisions can be taken today to create the best possible tomorrow. The process of foresight involves assessing the future implications of present actions, assessing the present implications of possible future events and considering desired future states. The use of foresight as a tool in policy and strategic decision making increased especially in the last decade of the twentieth century in order to enhance competitiveness and innovation of nations, regions, corporations and even individuals. However, it was observed that none of these definitions were capable enough to represent an integrated and holistic view about the impact of foresight on the management of the future. Human beings have been always trying to know the future but scientific study of the future is a recent phenomenon. Fifty years ago, the well-known French futurist, De Jouvenel, who is called the Father of Future Study, founded an institute, Fotorible, which is now considered as a powerful centre of future study in Europe. Later, he published the Futurist journal which literally means orienting to the future or building future (Khazaii, 2009). Futures have been known by several terms as Future Study, Futures, Future Field, Forecasting, Foresight, and Futurology which are commonly used in future studies. Future study is the systematic study of probable, possible, and preferable features and underlying attitudes, ideologies and myths of each future (Inayatullah, 2007). However, each of these terms underlines various theories and presuppositions and has their own particular methods. The publication of two books, The Image of the Future (Polak, 1961) and the Act of Conjecture (Jouvenel, 1967), represent modern views toward future study. Polak used the concept of image of the future to analyze the ups and downs of different civilizations. Jouvenel also codified a large number of principles of future study for the first time.

Foresight at national level first emerged in 1960 in defence sector of the U.S. First, no distinction was made between foresights and forecasting and it was done quantitatively. In addition to numerous organizations and activities which were contributing to future study, appropriate social structure (such as Jouvenel's science fiction stories, cultural confrontations (or conflicts), and the movement of environmental support in 1970s (Dater, 1979) facilitated its development. A few years later, since the bestselling book Alvin Toffler's Future Shock was published, interest in future has become fashionable. Since early 1970s, the science and art of foresight was formally applied as a tool for policy making in a limited number of countries especially Japan. Later, it entered the field of technology and civil. Since the beginning of 1980s, foresight has been differentiated from forecasting and undergone a shift from quantitative methods to networking (European Commission, 2006).

Scientometrics makes use of statistical and measurement methods to determine the development and growth criteria for science and its development level and influence on different human societies. This field first emerged in Soviet Union. In Eastern Europe and in Hungary in particular, it is used to measure sciences quantitatively at both national and international levels and for government and private institutes. Scientometrics is an interdisciplinary area which examines a wide range of topics due to its broad scope and is conceived of as a dynamic issue which deals with all quantitative aspects of sciences and scientific research. Quantitative evaluation of scientific activities as a significant developmental factor could help the authorities and planners to make the most use of financial and human resources and positively contribute to enhancing the socio-economic structures of the society. It is neither merely a diagnostic tool nor a magic panacea, but it is one of the best tools in clarifying the scientific issues and proposing viable solutions for various problems considering the research-related activities. According to Moraveik's theory, scientometrics is an interdisciplinary area which not only is related to the limited topic between two traditional problems but also entails numerous traditional topics due to its broad scope. Vinkler maintains that scientometrics is a science which deals with all quantitative aspects of science and scientific research.

Nowadays, with the advent of digital environment, all these areas have undertaken wide changes and scientometrics has come to the force in academic and scientific settings. Accordingly, in the past decade, the application of scientometrics has undergone wide, quantitative and qualitative changes. Nowadays, scientometrics is applied to ranking people, universities, research and scientific centres, countries, etc. Moreover, a large number of indices such as ISI, ISC, Scopus, etc. have emerged. Within the last recent years, foresight, as one of the main applications of scientometrics, has received considerable attention.

Both scientometrics and foresight are interdisciplinary and are quite similar in observing scientific activities, analyzing the trends, scientific prospective and predictions. For instance, when we deal with strategic view toward scientific issue, three steps should be taken: first, we should map the optimal condition for us i.e. in near future, where we should reach between the duration and far future which is characterized by foresight (futurology). Second, we should consider the current condition and level. Third, the gap between the current condition and the optimal one should be examined which is done through scientometrics tools and techniques and the road map is formed accordingly. Indeed, making use of foresight means taking action for making policies in order to realize an optimal future through which a number of other helpful results would be obtained: developing new networks and communications, goal-setting and enhancing the sense of shared commitment, publishing the information, and prioritizing the options and actions. In this article, the authors first present a brief historical overview of the formation of the term "foresight". Then, scientometrics and foresight are defined and the mutual relationship between these two terms in different scientific texts are examined and elaborated. Providing the related literature on scientometrics and foresight, the authors will discuss the status and role of scientometrics as a tool for foresight-related studies.

Foresight

The concept of foresight has been frequently changed since people hold different views based on their own area of knowledge and it has a short history and has undergone various experiences. This is why presenting a clear and precise definition of the concept seems difficult. Ben Martin, a pioneer in the field, presented the first generally accepted definition: "Foresight is a systematic attempt to look at long-term future of science, technology, environment, economics, and society, made to identify newly emerging technologies and enhance strategic research areas that seem to be of the most socio-economic benefits" (cited in Nazemi, 2006, p. 27). Luke Georghion maintains that it is a systematic tool for evaluating those scientific technological advances which can make a huge impact on individual competition, wealth creation, and life quality (Sahebinejad, 2006). Horton points to it as an extensive developmental process of attitudes regarding possible ways to develop future. Creating a complete understanding of these attitudes would lead to decisions which have the potential of creating the best possible future. Gavigan believes that it is a systematic and cooperative process entailing understandings of the future which provides a middle-term vision aiming at making up-to-date decisions and attracting shared actions" (Nazemi, 2006).

According to Bell (1996) the purposes of Futures Studies are to discover or invent, examine, evaluate, and propose possible, probable and preferable futures. Foresight is neither prophecy nor prediction. It does not aim to predict the future but to help us build it (European commission, 2006). In Webster dictionary, it is defined as "an organized and goal-oriented process which takes into account expectations of different actors about technology and codifies strategic vision of the future to support and confirm the extensive socio-economic development" (Nazemi, 2006). Loveridge (2009) views foresight as a description of a set of solutions for improving decision-building and decision-making methods in order to develop a strategic vision and intelligent forecasting. Richard Slaughter, the author of the book The Foresight Principle, also presented various definitions but the following definition is mostly cited: the ability to create and keep a practical, cohesive and qualitative vision, looking forward and taking advantage of understandings resulting from helpful organizational methods e.g. identifying improper conditions, guiding policy, the strategy of forming and examining new markets, products and services (Slaughter, 2007).

FOREN describes it as a systematic, cooperative process, intelligent community, and building middle-term and long-term vision which aims at forming decisions in order to set up the future activities (Miles, 2004). A futurology theoretician briefly defined it as knowing future and studying probable and desirable futures for a society (Mansouri, 1998). Accordingly, future is the reason behind the existence of the past and present and is a way for all to control their own future. In other words, the aim of vision-based thinking is clarifying the choices of our past and present in light of probable futures (Ghodeh, 1996). Joseph Kootez describes foresight as a process in which an individual strives to reach a more comprehensive understanding of the powers forming the long-term future which can be incorporated in formulating politics, planning and making decisions. Foresight entails both quantitative and qualitative instruments for studying growing signs and indicators. Being in direct relationship with analyzing political events is considered as its optimal state. It is also more useful than other fields. It prepares us for meeting the needs and future opportunities (Khazaee & ElahiDehaqi, 2013 cited in Monzavi, 2013).

Foresight tries to create a futuristic thinking style and understanding in business, governmental sectors and knowledge institutes in order to allow them to understand the probable opportunities and threats within the next 10 to 20 years in the realm of market and technologies. Then, it brings about and fosters cooperation among these three sectors and directs their activities to create competitive assets, to improve the quality of life, and to provide stable development (Sajjadipour, 2006). The main functions of foresight are as follows: orientation, recognizing newly developed trends, adapting objective to fulfil needs, supporting and promoting decisions and policies serving the preferences of the interested, promoting external relationships with those who benefit from the research or education, and determining priorities. (Khazaee & ElahiDehaqi, 2013). Table 1 presents a comparison of the definitions posed by various scholars. The commonalities and discrepancies of their definitions are highlighted in terms of eight aspects. Different definitions and viewpoints toward foresight indicate the wide scope of experts considering it. According to the indices presented in table 1, it may be inferred that being organized and systematic, constructing the vision and long-term future are the only commonalities of these definitions. Nevertheless, some scholars have pointed to it as a process and some others have considered it as unimportant.

Table 1. Comparison of the definitions posed by eight various aspects

Theoreticians	*Process*	*Being organized and systematic*	*Cooperative*	*Constructing the vision*	*Long-term future*	*Gathering the operations*	*Gathering perceptions*	*Making decisions*	*Making wise predictions*
Webster	*	*	*	*	*	*			
FOREN	*	*	*	*	*	*	*	*	
Martin	*	*	*	*	*				
Georghion		*	*	*	*				
Horton	*	*	*	*	*	*	*	*	
Gavigan	*	*	*	*	*	*	*	*	
Slaughter		*		*	*		*	*	*
Loveridge		*		*	*		*	*	*
Foundation of Future Development		*			*	*		*	
Kootez	*				*		*	*	
TagaviGilani & Ghofrani	*				*				*
Ghodeh				*					

Scientometrics

There is a wide consensus that the emergence of the term scientometrics coincided with coining the term book-metrics by Allan Perichard in Russian for the first time in 1969 which was used by Vasili, Namilov and Mulchenko in Soviet Union in order to study all aspects of existing written sources related to science and technology (Erar, 2002; Hood & Wilson, 2001; Wouters, 2003; Glanzel, 2003; Egghe, 1999; Schubert, 2002; Granovsky, 2001). The term Scientometrics is made up of two words: sciento which means science, and metrics which is derived from measuring (Ganji, 2004). With the publication of Scientometrics Journal in 1978 by Braun in Hungary, this field has witnessed many advances. Some of the works which contributed to the formation of the field of Scientometrics are as follows: Nalimov (1970), Nalimov and Mulchenko (1971), Nalimov *et al.* (1971).

Brooks (1999) asserted that the term was first emerged in a FID publication in 1969. The Scientometrics Journal was not that old but the term scientometrics was totally acknowledged in English (Garfield, 2007; Schubert, 2002). After the publication of the aforementioned journal, Braun established the scientometrics department in the Information and Scientometrics Research section of the library of Hungarian faculty of sciences. It was his attempts that led to the evolution of scientometrics and changed it to a wider area with more scientific communications in socio-economic studies.

There are many definitions for the term ”Scientometrics” in the literature. Scientometrics is the quantitative study of the disciplines of science based on published literature and communication. This could include identifying emerging areas of scientific research, examining the development of research over time, or geographic and organizational distributions of research (Glossary of Thompson, 2008). In a brief but general definition, Wilson (1999) defines scientometrics as the study of all quantitative aspects of science, related communications and scientific policies. Osareh cites Bookstein (1995) and provides a clear and brief definition of scientometrics and calls it “the knowledge of measuring science” (Osareh, 1997). Eom (2009) describes it as the application of quantitative instruments to study scientific communications. International encyclopaedia of Library and Information Sciences (2003) defines it as follows: “Scientometrics is a sub-discipline of Sociology of Science which is mostly related to scientific policies. It is concerned with quantitative study of scientific activities in general and publications in particular”. In an elaborate definition, Van Raan (1997) asserts that “scientometrics studies deal with quantitative study of science and technology which aims at knowledge development related to the developments of science and technology and questions considering scientific and social policies. The scientometrics studies are basically problem-oriented and interdisciplinary which take advantage of methods of Social and Behavioral Sciences, e.g. mathematical methods and statistics, models from Social Sciences, observation and survey from Psychology, computer sciences and so forth”.

Nalimov and Mulchenko, define it as “the application of quantitative methods which is concerned with analyzing science as an information process” (cited in Glanzel, 2003). Since Nalimov’s coinage of the Russian equivalent of the term ‘Scientometrics’ (naukometriya) in 1969, this term has grown in popularity and is used to describe the study of science: growth, structure, interrelationships and productivity (Hood & Wilson, 2001). San Gupta (1993) refers to the underling purpose of scientometrics as “evaluating the recent developments of any fundamental, scientific content and the effective factors in constant development of research activities in a particular area after World War II”. It is nowadays one of the most frequent methods for evaluating scientific activities and managing research. It entails quantitative examination of scientific products, scientific policies, and scientific communications, planning the scientific map of several areas of knowledge and drawing the science map. In scientometrics, scientific communications, methods of producing, distributing and taking advantage of scientific information are measured and evaluated. As a result, it is called the science of measuring science. It determines the achievements of a thinking area and predicts the probable lines for future developments through examining and exploring the underlying structure and system of a scientific area quantitatively. The underlying purpose of any activity in this field is providing the required information for technological, research, and scientific policy making and planning. This information, especially regarding scientific policy making at national level, has received considerable attention and enjoys several economic, social, political, and cultural dimensions (Osareh, 2009).

Mutual relationship between futurology and scientometrics

Nouroozi-chakoli maintains that "scientometrics and futurology were presented when the discussion of Science (with capital S) arose. Science gave rise to the emergence of these fields. In science (with small s), pursuing science was a personal tendency. Before World War II, science was more personal than under the pressure of organizations. In Science, organizations and even a country are in charge. Accordingly, it may disadvantage a country in the case of any loss. As a result, science generated a number of considerations one of which was future studies and science-studying (General Book of the Month, 2013).

One of the main principles of futurology is choosing a model and an appropriate method for gathering the data and information about several topics related to areas under the study and finding an efficient model and method to analyze the data in order to reach a better understanding of the future and facilitate the decision-making procedure (Abdollahkhani, 2011). There are various methods to depict future and foresight (Tabatabaiiyan, 2010). According to conducted research (World Future Society, 2004) there are over 33 methods to carry out futurology studies (Popper, 2008) see Figure 1.

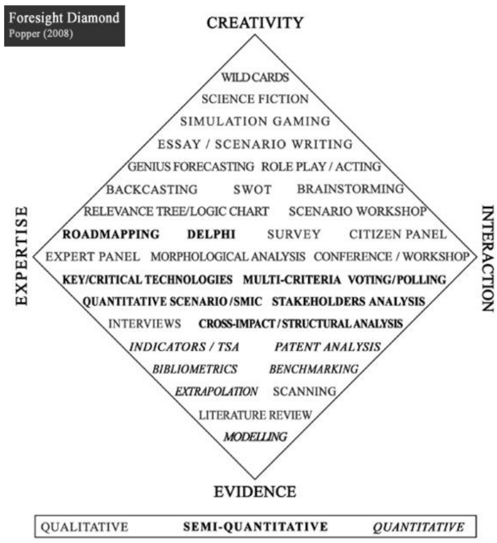


Figure 1. 33 methods to carry out futurology studies (Popper, 2008)

To clarify them more, Vinnari (2014) divides these methods into five categories in terms of the instruments and concepts used in futurology studies: 1. Methods for data collection (for example expert methods, Delphi, questionnaires), 2. Analysis methods (SWOT, trend analysis and cross impact analysis), 3. Tools for data organization (STEEPV and futures tables), 4. Tools for representing results (scenarios, back casting and futures images), and 5. Concepts for interpreting futures information (weak signals, megatrends and wild cards).

It should be noted that information is the key element in all of these methods. To analyze the obtained information, all methods make use of scientometrics instruments in one way or another. We have analyzed both the status quo and the future in scientometrics. It should be noted that both futurology studies and scientometrics studies are somehow interdisciplinary. Futurology has a special characteristic, i.e. being at large-scale which does not consider the details. However, it has a number of instruments; one of the most important instruments is scientometrics. Which road we should take and what we need for our purposes come from futurology studies which are achieved through small-scale studies in which scientometrics is involved. Scientometrics instruments can provide the futurologists with these types of information; otherwise, they would not be able to draw the appropriate map clearly (General Book of the Month, 2013). In recent years, many researchers have conducted scientometric analysis in different subject fields**.** For example (Osareh & Wilson (2002), Dutt, Garg & Bali (2003), Signore & Annovazzi (2004), Moin, Mahmoudi & Rezaei (2005), Wen *et al.* (2007), Mukherjee (2008), Tian, Wen & Hong (2008), Arruda *et al.* (2009)**.** A brief review of the related literature encompasses a number of evidence indicating the involvement of scientometrics in futurology. The available literature and comparing it with the nature of the studies (Georghiou and Keenan, 2008; Popper, 2008; FOREN network, 2001; Cassingena Harper, 2010; Calof and Smith, 2010; Smits and Kuhlmann, 2004; Havas *et al.*, 2010; Da Costa *et al.*, 2008) in the field of futurology indicate that the reports of scientometrics studies can be applied to futurology and contribute to drawing the road map of a country and achieving several visions regarding scientific, political, social, economical, and cultural issues. These studies significantly contribute to the growth and development of new technologies and inventions which the society requires (WMA, 2006). Moreover, the results of a simultaneous searching for the terms “futurology” and “scientometrics” or “Bibliometrics” in such valid information databases as Scopus, Science Direct, and Emerald. By the researchers suggests the growing increase in the number of published articles since 1990 till now which might point to more application of scientometrics instruments in futurology studies (Diagram 1).

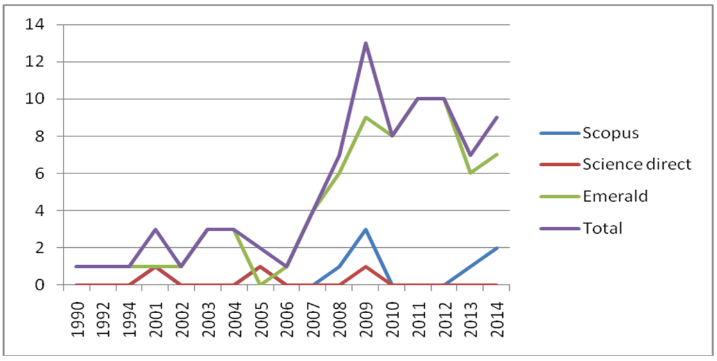


Diagram 1. The results of the terms “futurology” and “scientometrics” or “Bibliometrics” in Scopus, Science Direct, and Emerald since 1990 till 2014

The role of scientometrics and futurology in determining the vision and drawing the road map

The real independence and development of the countries is directly associated with their abilities in scientific production and research-scientific development. In other words, the development of the countries relies on predicting future and revising the performance and purposes, explaining the status quo and drawing the future development map in order to meet the local, regional, national and global needs. To this end, obtaining information regarding the scientific and research performance of the countries are essential. Acquiring this information is possible through conducting scientometrics studies (Fadaei & Hasanzadeh, 2010).

According to Kostoff *et al.* (1998) "The relationships among science and technology fields, and the temporal evolution of these relationships, have been of long-term interest to many organizations. The 'roadmaps' of these relationships have been used for science and technology marketing; science and technology management; enhancing communications among researchers, technologists, managers, users, and stakeholders; identifying gaps and opportunities in science and technology programs; technical intelligence; and identifying obstacles to rapid and low-cost product development. The generalized roadmap relates science and technology performed at some point in time to: its science and technology heritage; other relevant science and technology being performed at the same time; and future relevant science and technology and eventual end products".

Observing the scientific issues is included in topics of information sciences. Observation is employed in futurology and scientometrics. It is used to identify both the strengths and weaknesses and to enhance the former and to remove the latter ones. Indeed, it is of utmost significance in reaching the optimal status in scientific map of the country. In order to prepare the best road map, the issues should be observed in practice. It is considered as a key point in scientometrics and futurology in order to reach the optimal future. We have a vision for future which is a precise picture of what we strive to reach. The mutual understanding between future and us is called vision in which our goals are precisely determined. In fact, to reach an optimal status, detailed objectives, competitive goals, modelling, taking models, etc. should be precisely done in scientometrics and futurology (General Book of the Month, 2013). In order to determine the status quo and optimal status, scientometrics methods and content analysis are used. For instance, the number of recorded inventions, research and development expenses, the number of engineers and scientists, and the number of scientific articles in relevant areas, etc. are determined to achieve an understanding of the status quo and the optimal status (Tabatabaiiyan, 2010).

Conclusion

Scientometrics has received considerable attention as a useful instrument in decision-making areas in scientifically developed societies within the last decades (Georghiou and Keenan, 2008; Popper, 2008; FOREN network, 2001; Cassingena Harper, 2010; Calof and Smith, 2010; Smits and Kuhlmann, 2004; Havas *et al.*, 2010, Da Costa *et al.,* 2008). Nowadays, scientometrics instruments are used not only for contributing to large-scale decision-making and strategic areas but also for enhancing the quality of analytical data especially in universities (Ball & Tunger, 2004; Gorraiz *et al.,* 2010). The number of institutes which are established to meet the ends of scientometrics is growingly increasing (Gumpenberger, Wieland, & Gorraiz, 2012). In this article, the history and definitions of scientometrics and futurology were presented briefly. Moreover, the mutual relationship between these two fields was examined in various texts and their role in drawing the vision and road maps of the countries in different fields was elaborated. It was tried to deal with the role of scientometrics instruments and techniques in screening and determining the existing gap between the optimal status and the status quo and drawing the road map in futurology studies. It should be noted that scientometrics instruments are too limited to evaluate the complete status of an area in a system since the scientometrics instruments are focused on explicit knowledge and research and we have no instrument for implicit parts; what is done in that area cannot be observed since numerous scientific products are the result of unpublished researches. Having access to a well-defined vision requires operational, timely and integrative planning at different levels. This planning should take advantage of a set of resources, facilities, and talents so that it starts from status quo and moves toward the defined status in vision document (optimal status) through an organized and steady movement in a specified time. This is exactly futurology. Realizing this process requires drawing an accurate road map in which taking the road, predicting resources and facilities, dividing labour at national level and the way of participating and considerations are precisely and clearly specified. In this way, the vision is developed in detail and at smaller and more operational scales. This requires drawing a precise road map based on scientometrics instruments. Codifying the documents of scientific developments of the country and the comprehensive scientific map of the countries are in line with scientometrics studies. To put in a nutshell, futurology of any area should be based on accurate methodology and observation of the status quo through accurate scientometrics instruments and methods in order to see how many ways are available and to spot the problems. Moreover, metadata and instruments for analyzing them are needed in order to reach the minimum amount of information. Information sciences play a mediating role and set the required informational scenes. Through using appropriate information settings and using scientometrics instruments, accurate futurology in relevant areas would be possible in different countries. All in all, it is inferred that scientometrics and futurology topics are closely interwoven. Using scientometrics instruments in periodical screening of the road map and future road of various areas is both effective and inevitable.

References

Abdollahkhani, A. (2011). *Techniques of Forecast*. Tehran: Abrar Institute of Contemporary International Studies. [in Persian]

Arruda D., et. al. (2009). Brazilian computer science research: Gender and regional distribution. *Scientometrics*, 79 (3): 651-665.

Ball, R. & Tunger, D. (2004). Bibliometrische Analysen - ein neues Geschäftsfeld für Bibliotheken? *B.I.T. online*, 7(4): 271-278.

Bell, W. (1996). What do we mean by future studies?, in Slaughter R.(Ed) New thinking for a new, Millennium. London: Rutledge, 3-25.

Brooks, B.C. (1990). Biblio, sciento, infor-metrics? What are we talking about? In: L. Egghe, R. Roussean (eds), informetrics 89/90. *Selection of paper submission for the second international conference on bibliometrics, scientometrics and informetrics*, Amsterdam, Netherlands: Elsevier.

Calof, J. & Smith, J.E. (2010). Critical success factors for government-led foresight. *Science and Public Policy*, 37(1):31-40.

Cassingena Harper, J. (2010). European policy strategies and the role of foresight. paper presented at *ETEPS Seminar*, 27-28 September, Vienna.

Da Costa, O., Warnke, P., Cagnin, C. & Scapolo, F. (2008). The impact of foresight on policy making: insights from the FORLEARN mutual learning process. *Technology Analysis & Strategic Management*, 20(3):369-87.

Dator, J.A. (1979).The Futures of culture or cultures of the future. In A.J. Marsella, R. G. Tharp & T.J. Ciborowski (Eds). *Perspectives on Cross-Cultural Psychology*. New York: Academic Press.

Dutt, B., Garg, K.C. & Bali, A. (2003). Scientometrics of the international journal Scientometrics. *Scientometrics*,56(1):81-93.

Egghe, L. (1999). An explanation of the relation between the fraction of multinational publicationss and the fractional score of a country. *Scientometrics*, 45:291-310.

Eom, S.b. (2009). *An introduction to bibliometrics and informetrics*. Available at: http://www.igi-pub.com/downloads/excerpts/9671.pdf.

Erar, A. (2002). Bibliometrics or Informatics : Displaying Regulating in Scientific Patterns by using statistical Distributions. *Hacettepe Journal of Mathematics and Statistics*,31:113-125.

European Commission. (2006). *Using foresight to improve the science-policy relationship European Commission: Directorate-General for Research Directorate K –Social sciences and humanities*; Foresight Unit K2 – Scientific and technological foresight Scientific Officer.

Fadaei, G. & Hasanzadeh, H. (2010).Evaluation of scientific production in the field of the humanities faculty of the University of Tabriz(2005- 2007). *Journal of the Public Library of Iran*, 16(2):158. [In Persian]

FOREN Network (IPTS, PREST, CMI and SI). (2001). *A practical guide to regional foresight. Institute for Prospective Technological Studies (IPTS)*, Seville, available at: http://foresight.jrc.ec.europa.eu/ documents/eur8en.pdf (accessed 16 December 2011).

Foundation of Future Development. (2005). *Technology foresight methods*. Tehran: Foundation of Future Development. [In Persian]

Ganji, A. (2004). *Scientific Productivity of Academic Staff of University of Mashhad*. MS Thesis, School of Education, Ferdowsi University of Mashhad. [in Persian]

Garfield, E. (2007). From The Science of Science to Scientometrics: Visualizing the history of science with HistCite software. Presented at *11th ISSI Itnernational Conference, Madrid, June 25*. Retrieved from http://www.garfield.library.upenn.edu/papers/ issispain2007.pdf.

General Book of the Month. (2013). Report: *Scientometrics issues and future visions, maps, and map the scientific*,16(10): 93-99.[In Persian]

Georghiou, L. & Keenan, M. (2008). *Evaluation and impact of foresight*. in Georghiou L., Cassingena Harper J., Keenan M., Miles I. and Popper R. (Eds), The Handbook of Technology Foresight: Concepts and Practices, Edward Elgar, Cheltenham.

Ghodeh, M. (1996).from Predictions to strategic vision. Translated by M. salehpour. *Journal of the Plan and Budget*. No.11.

Glanzel, W. (2003). *Bibliometrics as a research field: a course on theory and application of bibliometric indicators*. [online]. Available at: http://www.nprslis.net/2004/bib_module_kul.pdf.

Glossary of Thompson scientific terminology. (2008). *The Thompson Corporation*. Retrieved from http://science.thomsonreuters.com/support/patents/patinf/terms/

Gorraiz, J., Gumpenberger, C. & Mayer, W. (2010). Bibliometric practices and activities at the University of Vienna. INFORUM 2010. *16th annual Conference on Professional Information Ressources*.

Granovsky, Y.V. (2001). Is it possible to measure science? V.V. Nalimovs research in scientometrics. *Scientometrics*, 52(2):127-150.

Gumpenberger, C., Wieland, M. & Gorraiz, J. (2012). Bibliometric practices and activities at the University of Vienna. *Library Management*, 33(3):174 – 183.

Havas, A., Schartinger, D. & Weber, M. (2010). The impact of foresight on innovation policy-making: recent experiences and future perspectives. *Research Evaluation*, 19(2):91-104.

Hood, W. & Wilson, C. (2001). The literature of bibliometrics, scientometrics, and informetrics. *Scientometrics*, 52 (2), 291–314.

Inayatullah, S. (2007). *Questioning the future: methods and tools for organizational and societal transformation*(3rd ed.). Taipei, Taiwan: Tarnkang university.

International Encyclopedia of Information and Library Scince. (2003). Edited by john Feather and Paul Sturges. New York : Routledge.

Jouvenel, B.D. (1967). *The art of conjecture*, Trans N. Lary, London: Weidenfeld and Nicolson.

Khazaee, S. & Elahi Dehaqi, I. (2013). Crucial Factors in the Efficacy of National Foresight. *Journal of Futures Studies*, 9(2):4-10. [In Persian]

Khazaii, A. (2009). *Necessary to allocate a larger share of GDP to research*. Fars News Agency. December 1. [In Persian]

Kostoff, R.N., Eberhart, H.J. & Toothman, D.R. (1998). Database Tomography for Technical Intelligence: A Roadmap of the Near-Earth Space Science and Technology Literature. *Information Messing & Management*, 34(I):69-85.

Loveridge, D. (2009). *Foresight: The Art and Science of Anticipating the Future*. New York and London, Routledge.

Mansouri, R. (1998). *Iran 2048: National commitment of scientific and cultural development*, Tehran: Tarheno publications. [In Persian]

Miles, I. (2004). *Three Worlds of Foresight EU-US Seminar: New Technology Foresight*, Forecasting & Assessment Methods-SEVILLE.

Moin, M., Mahmoudi, M. & Rezaei, N. (2005). Scientific output in Iran at the threshold of the 21st century. *Scientometrics*, 62:239–248.

Mukherjee, B. (2008). Scholarly literature from selected universities of Delhi and Uttar Pradesh: A pilot study. *LIBRES*, 18 (1). Retrieved from http://libres.curtin.edu.au/libres18n1/Mukherjee_Final_rev.pdf

Nalimov, V.V. & Mulchenko, Z. M. (1971). *Scientometrics. The study of science as an information process*, Nauka, Moscow, 1969. [In English-Measurement of Science. Study of the Development of Science as Information Process, Washington, DC: Translation Division, Foreign Technology Division, United States Air Force Systems Command, 13 October, 1971.]

Nalimov, V.V. & Mulchenko, Z. M. (1970). Science and biosphere: An attempt of comparison of the two systems, *Priroda*: 11: 55–63.[in Russian]

Nalimov, V.V., Kordon, I.V. & Korneyeva, A.YA. (1971). Geographical distribution of scientific information, In: *Information materials of the Academic Scientific Board on Cybernetics*, 2(49):3–37. [in Russian]

Nazemi, A. (2006). Regional foresight as land use planning Future studies *.conference technology and development prospects Tehran*, University of Technology(Amirkabir). [In Persian]

Osareh F., Wilson C.S. (2002). Collaboration in Iranian scientific publications. Libri, 52:88-98.

Osareh, F. (1997). A brief review of bibliometrics. *Quarterly book*. 33:90-108. [In Persian]

Osareh, F., etal. (2009). *The book surveys to web surveys: an analysis of the principles, attitudes, norms and criteria*. With an introduction by A. Horri. Tehran: Ketabdar. [In Persian]

Polak, F. (1961). *The image of the future*, (trans. E. Boulding), Leyden, Sythoff.

Popper, R. (2008). *Foresight methodology*. in Georghiou L., Harper J.C., Keenan M., Miles I. and Popper R. (Eds). The Handbook of Technology Foresight: Concepts and Practice, Edward Elgar Publishing, Northampton, MA, pp: 44-90.

Sahebinejad, M. (2006). *Technology foresight: a tool for priority setting of nanoscience and technology in Iran*. MS Thesis, Allameh Tabatabaeii university. [in Persian]

Sajjadpour, M. (2006). The role of values and ethics in Technology Foresight. *Imam Sadiq University Research Journal*, 24. [In Persian]

San Gupta, A.N. (1993). review of bibliometrics, informetrics, Scientometrics and measurement Library. Translator Vazirpour M., *information Science*,2(10):38-58. [In Persian]

Schubert, A. (2002). The Web of Scientometrics - A statistical overview of the first 50 volumes of the journal. *Scientometrics*,53(1):3–20.

Signore, A. & Annovazzi, A. (2004). Scientific production and impact of nuclear medicine in Europe: How do we publish?. *European Journal of Nuclear Medicine and Molecular Imaging*, 31(6):882-886.

Slaughter, R. (2007). *AFI Monograph Series*, vols. 1-10, Melbourne: Swinburne University. Available in pdf form from: http://www.swin.edu.au/agse/ courses/foresight/monographs.htm.

Smits, R. & Kuhlmann, S. (2004). The rise of systemic instruments in innovation policy. *International Journal of Foresight and Innovation Policy*, 1(1/2):4-32.

Tabatabaiiyan, H. (2010). *Evaluating the existing and desired technology in Iran*. Tehran: Center for Academic Publication. [In Persian]

Tian, Y., Wen, C., Hong, S. (2008). Global scientific production on GIS research by bibliometric analysis from 1997 to 2006. *Journal of Informetrics*, 2:65–74.

Van Raan, A.F.J. (1997). Scientometrics: state of the art. *Scientometrics*,38(1):205-218.

Vinnari, M. (2014). *An Introduction to Future Studies Approaches and Foresight*. LYY-course, University of Eastern Finland.

Wen & et. al. (2007). Scientific production of electronic health record research, 1991–2005. *Computer Methods and Programs in Biomedicine*, 86:191–196.

Williamet Management Associate(WMA). (2006). *Estimating intelectual property remining useful life*. In intelectual property analysis. White papers. Available at: http://www. Intelectualpropertyanalysis.com/article2.html

Wilson, C.S. (1999). Information annual review of information science and technology. *ARIST*, 34:107-247.

World Future Society. (2004). *The Art of Foresight*. A Special Report from the World Future Society. World Future Society.

Wouters, P. (2003). Nerdi Web on Science and the Internet. *Dlib Magazine*,9:3.

Indicators to Measure Genetics Literature: A Comparative Study of Selected Countries

S. L. Sangam*, Devika Madalli** and Uma Patil*

*Department of Library & Information Science, Karnatak University Dharwad-580 003, India
slsangam@gmail.com

**DRTC, Indian Statistical Institute, Bangalore -560059, India
devika@drtc.isibang.ac.in

Introduction

Indicator is a device providing specific information on the state or level of something. It is synonyms to measure, standard, yardstick, guideline, and test. These are reliable indicator of performance. Scientific performance is essentially a multidimensional concept, which cannot be measured by a single universal indicator. There may be a number of imperfect or 'partial' indicators, each representing a different aspect of research performance, with varying degree of success. Nonetheless, publications in the refereed scientific journals constitute the most important indicator of research performance of human life. In the present study an attempt has been made to study the Genetics literature.Knowledge of genetics being basic to progress in biology, agriculture, medicine, biotechnology, forensic sciences and many other fields, results of such studies are found highly useful. There are several dimensions of national science indicators can be used to study different aspects of the research output in the field of Science and Technology viz: Jain & Garg,(1992), Garg & Dutt (1999) Garg & Padhi (1999) Garg, Kumar & Dutta (2011) Sangam, (2002,2009,2010) , etc.

Methodology

For the present study the data are collected from PUBMED. It is a free Search engine primarily the MEDLINE database of references and abstracts on life sciences and biomedical topics. The United States National Library of Medicine (NLM) at the National Institute of Health maintains the database as part of the Entrez System of information retrieval. As of 18 April 2014 Pubmed has over 23 million records going back to 1966. Several Scientometric indicators have been suggested in the literature to measure national performance namely Growth, Doubling time, Activity Index, Attractivity Index, priority index, Impact Factor, H-index etc. will be used wherever necessary. The study compares Indian research priorities of 16 sub-specialties of genetics with 13 countries over the period of 20 years (1993 to 2012), in two block period 1993-2002 and 2003-2012.

Growth pattern of World and India

The total publication output of the world and India has been shown in Table-1 along with the growth rate and doubling time for over a period of 20 years (1993-2012). The table shows that the relative growth rate of world publications output decreases gradually from 0.735 to 0.082, but seems to be constant from 1999 to 2006 and from 2007 to 2012. Correspondingly the doubling time increases from 0.942 to 8.45 during 1993 to 2012. The mean growth rate and doubling time for the world is 0.19 and 5.13 respectively. The Indian output also decreases gradually from 0.779 to 0.152 except during 2000-03 (0.223-0.214) and 2012 (0.152). The doubling time correspondingly increases from 0.889 to 4.548 with a decrease during 2000-03(3.108-3.233). The average growth rate and doubling time for India is 0.25 and 3.31.

Table 1: World vs. India Growth Rate & Doubling Time

Year	World	Cummulative	GR	D(t)	India	Cummulative	GR	D(t)
1993	47334	47334			177	177		
1994	51408	98742	0.735282	0.942496	209	386	0.779688	0.888817
1995	56244	154986	0.450824	1.537184	254	640	0.505631	1.370565
1996	58939	213925	0.322291	2.150233	318	958	0.40338	1.717985
1997	62913	276838	0.257807	2.688057	325	1283	0.292109	2.372405
1998	70035	346873	0.225526	3.072813	352	1635	0.242442	2.858419
1999	76257	423130	0.198721	**3.487306**	363	1998	0.200504	**3.456292**
2000	83711	506841	0.180518	3.838955	499	2497	0.222943	3.108413
2001	87200	594041	0.158751	4.365327	610	3107	0.218568	3.170643
2002	88872	682913	0.139419	4.970623	758	3865	0.218304	3.174472
2003	94640	777553	0.129784	5.339627	924	4789	0.21436	3.23288
2004	100154	877707	0.121161	5.719662	998	5787	0.189292	3.661003
2005	104231	981938	0.112215	6.175626	1189	6976	0.186862	3.708626
2006	108823	1090761	0.105103	6.593549	1403	8379	0.183253	3.781659
2007	113615	1204376	0.099086	6.993926	1577	9956	0.172447	4.01863
2008	118540	1322916	0.093877	7.382016	1798	11754	0.166018	4.174241
2009	125142	1448058	0.090385	7.667205	2089	13843	0.163586	4.236302
2010	132054	1580112	0.087272	7.940656	2275	16118	0.152157	4.554507
2011	139994	1720106	0.08489	8.163488	2515	18633	0.144998	4.779391
2012	147004	1867110	0.082006	8.450615	3067	21700	0.152378	4.547899

Application of Growth Models for Genetics in India & World

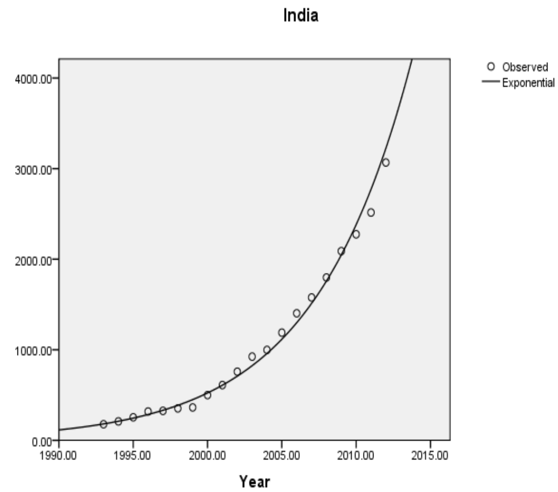


Fig. 1

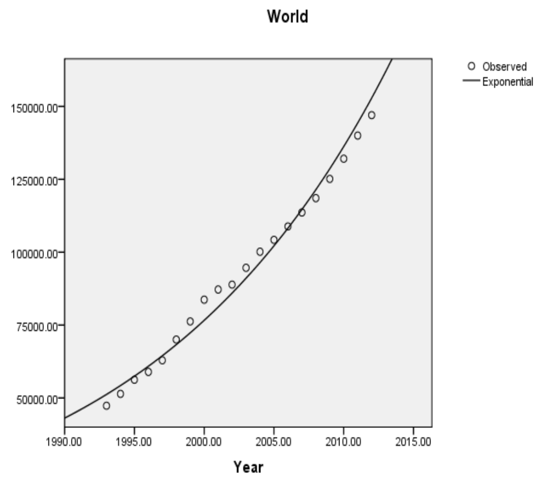


Fig. 2

Figure 1 & 2: Exponential growth curve for India and the World from 1993 - 2012

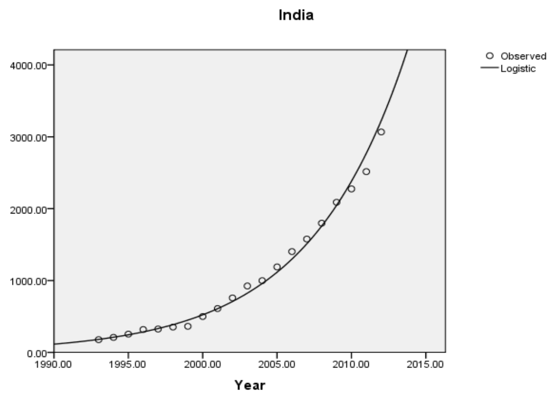


Fig. 3

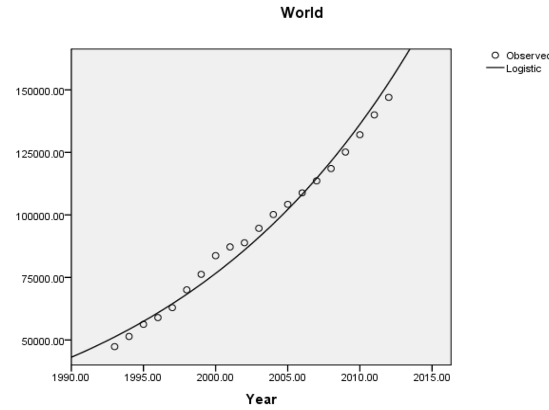


Fig. 4

Figure 3 & 4: Logistic growth curve for India and the World from 1993 - 2012

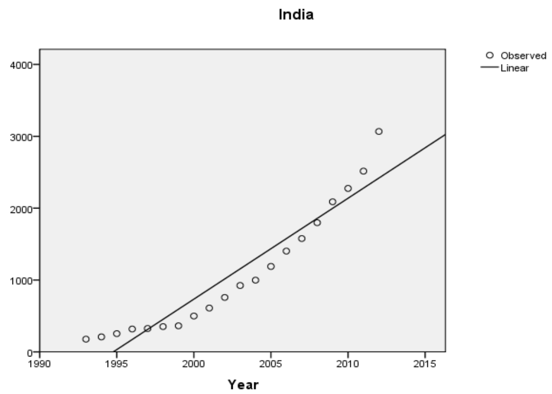


Fig. 5.

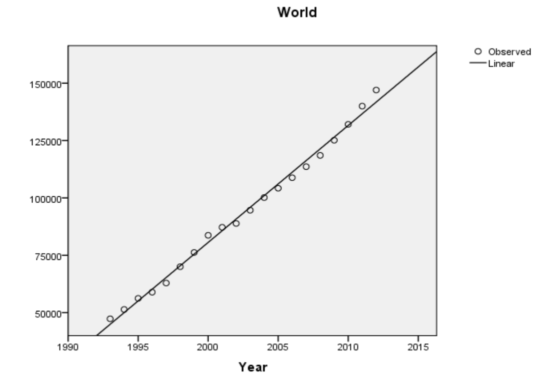


Fig. 6.

Figure 5 & 6: Linear growth curve for India and the World from 1993 - 2012

In order to get a clue in the selection of the best growth model, the plots of growth rate functions for different mathematical models are presented and visualized as in the Growth Models. From the table it is clear that for India's publication data the two growth models viz. Exponential and Logistic models (99.2%) are equally appropriate growth models. Whereas, for World's publications data Logarithmic and Linear growth model fits well (99.4%) compared to other growth models. Also the graphs for all these models for both world and India are presented. Overall observation in the application of Growth Models for the genetics publications exist that the world output best fits compared to the Indian output.

Activity Index

Activity index was proposed by Schubert and Braun (1986). It characterizes the relative research effort a nation or an institution devotes to a given subject field or sub-field and takes into consideration the effect of the size of the country as well as the size of the sub-specialty.

The value of AI=100 indicates that the research effort of a country/institution in a given field corresponds precisely to the world's average; AI >100 reflects higher than average activity and AI <100 lower than average effort dedicated to the field. The major advantage of using activity index over raw (absolute) count of publications is that it takes into account both the size of the nation/institution as well as the size of the discipline.

This has been demonstrated using global output in the field of genetics for the period 1993-2012. The data on the publication output and activity index of twenty major countries in ten sub-specialties of genetics. The data (Table-2 not given due to short place) has been collected using the advanced search option of PubMed. It is observed that USA tops the list (48%). This is followed by Japan; Germany and China. These four countries together produced about 74% of the total output. From this it can be inferred that different countries emphasize on different sub-specialties in the field of genetics.

Priority Index

Publication performance of Major countries

The distribution of publications in major countries in different sub-specialities of genetics for two block periods 1993-2002 and 2003-2012 is presented in Table-3. The publications in twenty different countries are arranged in the form of a matrix where the rows represent the countries and columns the sub-specialties.

Table 3: Branch wise publication output of major countries in Genetics (1993-2002)

	DG	EV	GE	G	HG	MG	MiG	MoG	PG	QG	Total
USA	15083	4038	24021	4792	110593	50378	4284	125253	11622	6217	356281
Germany	2118	746	3409	577	16785	3358	797	24164	1685	1221	54860
France	1640	583	2639	467	11865	442	911	16616	1855	1073	38091
Canada	1445	416	2197	423	9443	2761	480	12520	1325	595	31605
Italy	639	322	1438	211	8576	1254	371	8589	1439	527	23366
Australia	616	304	906	313	6312	4927	275	7725	1094	357	22829
Brazil	61	69	95	29	806	166	98	1180	361	58	2923
Mexico	100	58	186	48	583	65	49	1396	203	51	2739
Argentina	45	22	74	2	317	11	34	638	159	25	1327
S Africa	33	36	57	10	618	430	74	628	155	7	2048
Japan	2756	644	4907	678	31087	13320	1322	33869	2397	1391	92371
China	193	89	695	74	3667	2241	75	3337	494	207	11072
India	116	58	489	62	1387	791	201	2155	394	62	5715
Taiwan	99	53	350	25	2349	1293	100	2230	355	96	6950
Israel	366	101	621	88	4310	2951	98	3544	481	156	12716
Georgia	150	108	280	38	1580	844	157	2546	277	74	6054
Turkey	19	2	50	0	581	417	23	245	137	11	1485
S Korea	52	9	164	3	479	122	32	662	49	19	1591
Russia	41	71	146	24	536	204	42	1366	139	27	2596
Hong kong	65	8	78	13	738	39	36	703	149	75	1904
Total	25637	7737	42802	7877	212612	86014	9459	249366	24770	12249	678523

The data in table-3 covers articles published in different branches of genetics during the block period 1993-2002. Columns represent 10 branches and the rows 20 different countries. The data reveals that major contribution is from USA (3,56,281) followed by Japan (92,371); Germany (54,860); France (38,091); Canada (31,605); Italy (23,366); Australia (22,829); Israel (12,716); China (11,072); Taiwan (6,950); Georgia (6,054); India (5,715); Brazil (2,923); Mexico (2,739); Russia (2,596); South Africa (2,048); Hong Kong (1,904); South Korea (1,591); Turkey (1,485) and Argentina (1,327).

Maximum number of articles published during this block period among the branches of genetics is from Molecular Genetics (2,49,366) followed by Human Genetics (2,12,612); Medical Genetics (86,014); Genetic Engineering (42,802); Developmental Genetics (25,637); Population Genetics (24,770); Quantitative Genetics (12,249); Microbial Genetics (9,459); Genomics (7,877); and Evolutionary Genetics (7,737).Which indicates the importance of the branch Molecular Genetics, as molecular approach is needed for the final conclusion in all the fields of biological sciences.

The data in table 10 (Table not given) reveals articles published in 10 branches, represented in different columns and 20 countries in rows during 2003-2012. In the present block period also, it is USA which has maximum number of articles published (5,69,841) however, it is followed by China (1,36,427); Japan (1,25,929); Germany (97,530); Canada (52,484); France (52,224); Italy (50,664); Australia (35,580); India (29,140); Taiwan (23,347); Israel (21,794); Brazil (16,124); Georgia (9,248); Turkey (7,888); Mexico (7,373); South Korea (8,751); Hong Kong (6,131); Russia (5,860); Argentina (4, 641) and South Africa (4,440).

Relative Priority of sub-specialties of Genetics in different countries

Using the values of PI, we can identify the priorities and potential holes in the research agenda of 20 countries. For this purpose, we use a 7-point scale for qualitative description of research priorities, suggested by Nagpaul and Sharma (1999).

7 Point Scale for qualitative description of research priorities,

Scale	**Values**	**Description**
PI<25	1	Field of neglect
25<PI<55	2	Field of very low priority
55<PI<85	3	Field of low priority
85<PI<115	4	Field around the mean position of the country
115<PI<145	5	Field of marginal priority
145<PI<175	6	Field of high priority
PI>175	7	Field of thrust

Table-4 (Table not given) depicts the subject wise priority by 20 different countries with respect to 10 sub-specialities of genetics during two block periods 1993 – 2002 and 2003 - 2012. The priorities have been set in to seven point scale for the analysis of the priority.

Thrust Priority areas

Brazil has given maximum importance to the study of Environmental, Microbial and Population genetics during the first block period 1993-2002 and are treated as the thrust area of research and during the second block period 2003-2012 the thrust priority was given to population genetics excluding Environmental and Microbial genetics. Mexico has given thrust priority to EV and PG during both the block periods. Argentina-MiG and PG during both block periods. S. Africa-MiG and PG during first block period and EV, MiG and PG during second block period. India’s top priority during 1993-2002 was for MiG and PG. However no subfields were on thrust priority during the second block period. Israel has exhibited the thrust priority for MG during both block periods which indicates the health consciousness of the people of Israel. Georgia has treated MiG as its thrust priority during 1993-2002 and no thrust priority areas during 2003-2012. MG and PG were treated as thrust priority areas during both the block periods by Turkey. Russia has paid its thrust priority during first block period to EV and no thrust preference areas of research during second block periods. Hong Kong had thrust priority for PG, QG and only QG during the first and second block periods respectively.

High Priority areas

France has paid high priority for research in subfields MiG, QG and EV during the first and second block periods respectively. Italy has high priority for PG during first block period and no subfields were on high priority during second block period. Australia considered Medical genetics on high priority during 1993-2002 block period and no subfield is on high priority during second block period. Brazil has no subfields on high priority during first block period and MiG has been put on high priority during second block period. Mexico paid high priority for research on genomics in first block period and no subfield has put on high priority during second block period.

Marginal priority areas

Marginal priority was given by USA to studies on genomics only, during the first block period (1993-2002). Germany considered EV, MoG and QG as areas with marginal priority in both block periods accept MoG in the second block period. EV, MoG and PG were the subfields with marginal priority during first block period by France. However, the same position was occupied by DG, MiG, PG and QG during second block period. Studies on DG, EV, G and PG subfields were on marginal priority accept PG in the second block period in Canada. Italy has EV, HG and QG on marginal priority during the first block period and in the second block period the position was occupied by HG and PG. Australia had studies on EV, G and PG on marginal priority during first block period and EV, G, MiG and PG during second block period. Publications on EV and MoG were on marginal priority during 2002-2012 in Brazil.

Average Priority Areas

In the seven point scale, average priority area falls between Marginal and low priority areas. The highly developed countries like USA, Germany, France, Canada, Italy, Australia, Japan etc have considered many of the subfields with average priority. USA considered all the ten subfields of genetics with the equal importance evidenced through their inclusion in thye average priority areas in both the block periods except Genetics which falls in marginal priority area only during the first block period. This indicates the average preference of USA for all the subfields. Germany treated research on DG, GE, G, HG, MiG and DG, GE, G, HG, MiG, MoG as average priority areas for both the block periods. France has average preference for DG, GE, G, HG and GE, G, HG, and MoG respectively for both block periods. Canada had average preference for GE, HG, MiG, MoG, QG for the first block period and GR, HG, MiG, MoG, PG, QG during the second block period.

Low Priority Areas

Well developed countries like USA and France, have not considered any of the genetics subfields under low priority areas for both the block periods. Germany in first block period had only PG and during second block period MG and PG in low priority area. Canada has publications in the subfield of medical genetics under low priority area for both the block periods. However, Italy considered DG, G and only DG under low priority for both block periods respectively. Australia had low priority for DG and GE during first and GE and MG during second block period. DG remained as low priority area for both block periods for Brazil and GE, G were added to the second block period along with DG. Mexico has low priority for HG during first block period and for second block period it was DG, G, HG, and QG. Argentina considered HG and HG, QG respectively for both block periods as low priority areas. However, S.Africa has Mo G on low priority for first and GE, HG, MG, QG second block period. Japan being a developed country also had DG, EV, GE, G, PG, QG as low priority areas in first and EV, G, PG in second block period. China had EV, G, MoG and India had HG and QG in low priority areas of research during first block period however, DG, EV and EV, HG, QG during second block period. During 1993 – 2002, Taiwan has considered EV, GE and QG as low priority area during first block period and only MoG during second block period.

Very Low Priority Areas

Very low priority areas indicate the genetics subfields with the lowest priority which is indicated by the number of publications in the respective subfields. Germany considered medical genetics as the very low priority research area during the first block period. Medical genetics was in the very low priority during both block periods for Italy. Brazil treated the MG as low priority areas in both block periods along with GE during second block period. Mexico also considered the same subfield under low priority area during second block period. Argentina during second block period considered G and MG as very low priority area. South Africa puts three subfields viz., DG, GE, and G as very low priority areas of research during first and only DG in the second block period. DG and MiG subfields were in the very low priority during first block period in China. However, India has treated DG as the VLP area during both block periods. For Taiwan, it was DG, G during first and DG, EV for second block periods at VLP.

Neglected areas

Neglected sub-fields of genetics by different countries includes MG in both block periods by France; publications on MG during 1993-2002 by Mexico; Argentina also during first block period treated research on MG alone with G as neglected field. S. Africa has considered publications on QG during first block period as neglected area while Turkey on EV and G during both block periods. S. Korea treated G in the first block period as neglected area. Publications on MG were on the neglected part of studies in both block periods by Hong Kong.

Citations per Paper

Citation per paper is the most widely used indicator in bibliometric studies. It is a relative indicator computed as the average number of citations per publication. It normalizes the wide disparity in volume of literature published by prolific publishing nations and other smaller nations for a meaningful comparison of research influence. It is the ratio of total number of citations to the total number of publications. In case, where citations are not available, one can use normalized impact per paper.

The Table 5 (Table not given) shows the data on total number of publications and citations received in the field of genetics research during 1996 to 2012 has been collected from Scimago Journal Rank (SJR) indicator both for India and world arranged year-wise. From the table it is observed that in case of world, the CPP value has increased from 30.40 to 36.28 during 1996 to 2000 followed by decrease in CPP from 36.28 to 0.75 during 2000 to 2012. In India also, the CPP has increased till 2002 and has decreased for 18.35 to 0.48 during 2002 to 2012. This decrease in CPP may be because of lack of time lag by the recent publications to receive more citations.

Authorship and Collaboration

The present era is witnessing the practice of collaboration which is spreading very fast owing to the globalization of information. The days of individual research are gone. The present situation compels on the researchers to go for collaboration in research, thus resulting in the shift from solo research to team research.

Table 6: Publications by number of authors in two block periods

Ten years Block	Single Author	Two Author	Three Author	Four & above authors
1993-2002	191 (5%)	925 (24%)	936 (24%)	1811(47%)
2003-2012	433 (2.5%)	2501 (15%)	3184 (19%)	10846 (64%)
Total	**624 (3%)**	**3426 (16.5%)**	**4120 (20%)**	**12657 (60.7%)**

The table 6 reveals the following growth rate during two block periods in all the categories of co-authorship publications. The proportion of single-author publications has decreased from 5% during 1993-2002 to 2.5% during 2003-2012 with the average percentage being 3% for the entire period. The proportion of two-author publications has also decreased from 24% during 1993-2002 to 15% during 2003-2012 with the average percentage being 16.5% for the entire period.

The proportion of three-author publications has again decreased from 24% during 1993-2002 to 19% during 2003-2012 with the average percentage being 20% for the entire period. Whereas the proportion of four and more than four-author publications has increased from 47% during 1993-2002 to 64% during 2003-2012 with the average percentage being 60.7% for the entire period.

Highly productive authors in India

The table 7 (Table not given) shows 20 highly productive authors in the field of Genetics during 1993-2012. These authors together have contributed 3348 articles. Author Kumar S. is the highly productive author with 329 papers, followed by Kumar A. with 310 papers, Kumar R. with 214 papers, Singh S. with 200 papers, Sharma, S. with 192 papers, Sharma, A. with 180, Mittal, B. with 178 papers, Ghosh, S with 168 papers, Singh, L. with 162 papers, and Gupta, S. occupies 10th position contributing 160 papers.

Conclusion

By using different Scientometric indicators like activity index and attractivity index it can be identified whether the country is doing more or less research in a particular field or sub-field as compared to other nations. Is it doing better than others? Is it doing more research in a particular field compared to some other field, or is it doing better in one field compared to another. It can identify topics with significant increase in world publication output (*hot topics*); topics with significant decrease (*cold topics*); and topics with no significant increase or decrease in world publication output (*stable topics*). If a country publishes much less than the world average on a hot topic, it implies that the country has failed to pick up new developments and it needs some exploration. For stable topics, equal or above world average activity is a sign of healthy development, while a significant lower activity indicates a weakness. In the present study all the major countries have a mixed profile.

References

Garg, K. C., Kumar, S., & Dutt, B. (2011). Simple technique to normalize impact factor of journals, *DESIDOC Journal of Library and Information Technology* 31(5), 371-376.

Garg, K.C., & Padhi, P. (1999). Scientometrics of laser research literature as viewed through the Journal of Current Laser Abstracts *Scientometrics, 45*, 251-268.

Garg, K.C., & Dutt, B.(1999). Geographical distribution of the Indian science activity (in) *Emerging trends in scientometrics* edited by Nagpaul, P.S., Garg, K.C., Gupta, B.M. *et al*., Allied Publishers LTD, New Delhi.

Jain, A., & Garg, K. C. (1992). Laser research in India: Scientometric study and model projections, Scientometrics 23(3), 395-415.

Sangam, S. L., Liang, L., & Ganjihal, G. A. (2010). Modeling the Growth of Indian and Chainese Liquid Crystals Literature as reflected in Science Citation index (1997-2006). *Scientometrics*, *84*(1), 49-52. (DOI 10.1007/s11192-009-0079-x)

Sangam, S. L., Keshava, & Agadi, A. B. (2009). Growth Pattern of Marine Engineering Literature, *Information studies*, *16*(2), 113-120.

Sangam, S. L., & Gupta, B. M. (2002). Modeling the growth of world social science literature, *Scientometrics*, 53(1), 161-164.

Schubert, A., & Braun, T. (1986). Relative indicators and relational charts for comparative assessment of publication output and citation impact. *Scientometrics, 9*, 281- 291.

Analysis of the Financial Assistance to Non-ICMR Biomedical Scientists by Indian Council of Medical Research (ICMR) 2009 - 2013

Sandhya Diwakar and K. K. Singh

Indian Council of Medical Research, Ansari Nagar, New Delhi - 110029 (India)
sandhyadiwakar@gmail.com, singhkeshari@yahoo.com

The Indian Council of Medical Research (ICMR) has tried to address the issues relating to medical research and was also interested in capacity building at all levels for medical research, in partnership with international health organizations. The Council promotes Biomedical Research in the country through its several Human Resource Development programmes. This paper gives a study of financial assistance to Non-ICMR Bio-medical scientists and young scientists for participating in international scientific event, training programmes and short-term workshops/courses. Topic covers Science Policy and Research Evaluation.

Health research is the key to a well-functioning and effective health sector in the country. Major scientific breakthroughs hold the promise for more effective prevention, management and treatment for an array of critical health problems. The research to be undertaken should be on country specific health problems essential for the formulation of sound policies and plans for field action. But new interventions and development of new health products (drugs, diagnostics and vaccines) are possible only when there is well defined funding, infrastructure and priority for health research. Medical research in the country needs to be focused on new therapeutic drugs/vaccines for tropical diseases, normally neglected by multinational pharmaceutical companies on account of their limited profitability potential. In addition, India is also witnessing the ‘dual disease burden’ with the non-communicable diseases like cardiovascular diseases, diabetes, cancers etc. threatening to overtake infections. In the Government sector, such research has been confined to the research institutions under the Indian Council of Medical Research, and other institutions funded by the Central/ State Governments.

Since its establishment, the ICMR has been making concerted efforts to address the health needs of the nation. Given its limited resources – human, financial and infrastructural the Council has discharged its national obligations through its network of 31 national institutes including Six regional medical research centres, over 100 field stations and a strong and vibrant extramural research in medical colleges and other institutes.

To provide an opportunity to academic scientists and trainees and to provide a stimulus for those working or contemplating working in the field of medical science, this program provides international travel grants that can be used to acquire new research techniques and to promote collaborations. The applicants should be Bio-medical scientist engaged in R&D work. Senior Scientists (above 35 yrs of age) working in academic institutions and research laboratories and young scientists (below 35 yrs of age) including medical graduates, post-graduates and research scholars are eligible to apply to international scientific events.

From 2009 through 2013, ICMR ran a travel grants program that enabled Indian scientists to participate in international conferences, seminars, workshops and symposiums. The amount sanctioned for the program was INR 3.5 Crores (about USD 600), all he amount was disbursed. During that period 4274 travel grant applications were received out of which 1740 applications were approved for funding and 998 applicants finally availed the grant. Travel grants went to individuals at many institutions in the country and provided support for a wide range of biomedical research activities. An outcomes survey can be conducted to enhance the overall value and utility of the travel grant program.

Methodology

Data for the four year period during 2009-2013 was collected. Data points included name of the scientist, institution, designation, age, gender, state, conference title, venue, area of medical science, amount sanctioned/released and whether the application was approved, availed or rejected. The collated data was studied to identify the distribution of applications by country/state, area of medicine, designation; institution etc and inferences have been drawn from the study.

Observations

Table 1

State	*A*	*NA*	*Approved*	*R*	*Grand Total*
New Delhi	361	220	582	738	1320
Karnataka	117	71	188	265	453
Haryana	94	51	145	180	325
UP	89	99	188	273	461
Tamil Nadu	69	51	120	193	313
Maharashtra	68	48	116	247	363
WB	50	34	84	105	189
Punjab	42	48	90	119	209
AP	28	26	54	103	157
Gujarat	22	18	40	95	135
Kerala	18	16	34	37	71
Rajasthan	8	9	17	37	54
MP	7	13	20	39	59
HP	6	8	14	8	22
Assam	5	2	7	2	9
Odisha	5	7	12	23	35
J&K	4	2	6	14	20
Chhattisgarh	2	5	7	17	24
Uttarakhand	2	0	2	9	11
Meghalaya	1	4	5	3	8
Tripura	0	0	0	3	3
Goa	0	1	1	4	5
Jharkhand	0	2	2	4	6
Sikkim	0	1	1	2	3
Puducherry	0	3	3	7	10
Bihar	0	3	3	6	9
Grand Total	998	742	1740	2534	4274

New Delhi led all other states in terms of the applications submitted with a maximum number of 1320. Uttar Pradesh ranked second with the 461 applications submitted followed by Karnataka (453), Maharashtra (363), Union territory of Chandigarh (325), Tamil Nadu (313), Punjab (209), West Bengal (189), Andhra Pradesh (157) & Gujarat (135) These states accounted for three-fourth of the total applications received by ICMR. A zone-wise analysis indicates that North Zone is the most active with highest number of applications for grants received, approved and availed.

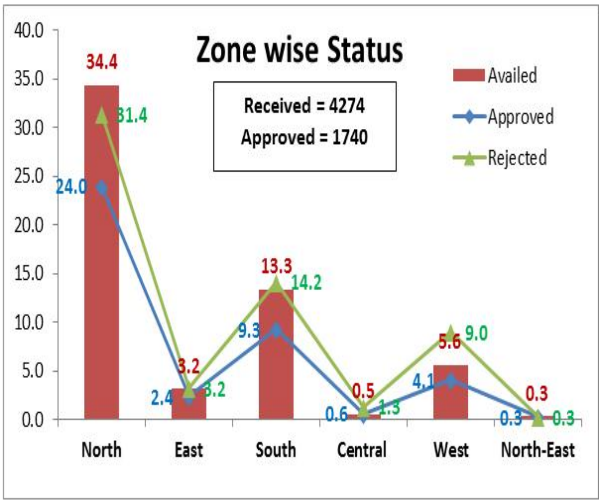


Fig 1. Zone wise distribution of applications received, approved, availed

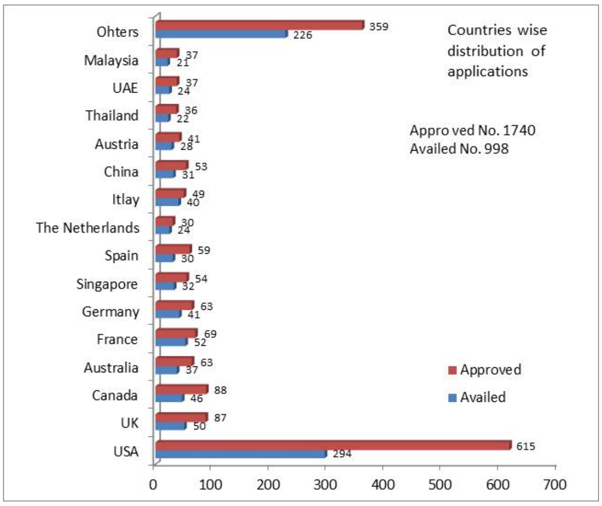
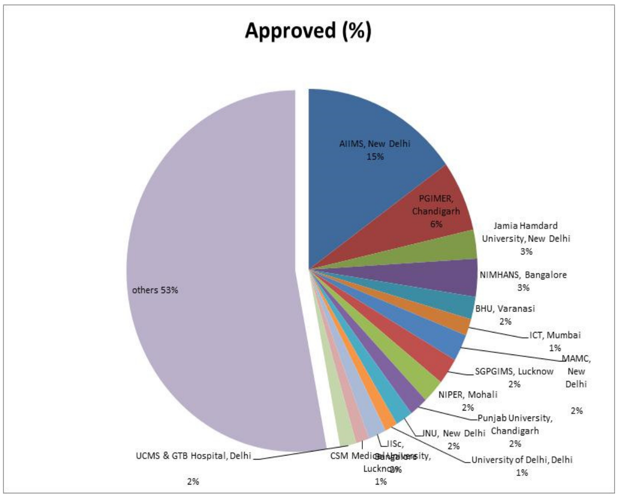


Fig 2. Countires wise distribution of applications approved, availed Top 15.

USA led all other countries in terms of the applications submitted with a maximum number of 1540 approved (615), UK 208 approved (87), Canada 204 approved (88), Australia 162 approved (63), France 161 approved (69), Germany 141 approved (63), Singapore 129 approved (54), Spain 129 approved (59), The Netherlands 125 approved (30), Italy 124 approved (49), China 122 approved (53), Austria 120 approved (41), Thailand 99 approved (36), UAE 92 approved (37), Malaysia 91 approved (37).



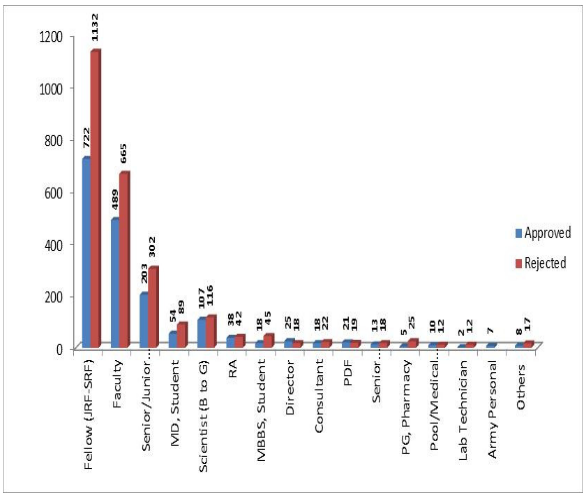


Fig 3. Share of Top 15 institutes which availed grants, n = 1740

Fig 4. Designation-wise breakup of approved applications n = 1740

In terms of the bio-medical science discipline, it was noted that the applications were received in wide range of areas such as Pharmaceutical Sciences - 87, Cancer/Radiation – 76, Neurosciences – 68, Vision and Ophthalmology – 51, Nuclear Medicine – 36, Drug Discovery

& Therapy – 55, Paediatrics - 37, Infectious & Emerging Infectious Diseases – 30, Human Genetics – 24, Psychiatrics – 28, Respiratory Diseases – 26, Immunology – 23, Environmental Sciences and Engineering – 16, Others - 1183 so on. There were 300 areas under which scientists had submitted applications. Out of these, the top 15 areas where highest number of applications received have been shown in Table 2.

Table 2.

Field of Research	Approved	Rejected	Total
Pharmaceutical Sciences	87	197	284
Cancer/Radiation	76	104	180
Neurosciences	68	88	156
Vision and Ophthalmology	51	69	120
Nuclear Medicine	36	63	99
Drug Discovery & Therapy	55	99	154
Paediatrics	37	50	87
Inf. & Emerging Inf. Diseases	30	45	75
Human Genetics	24	50	74
Psychiatrics	28	35	63
Chest/Respiratory Diseases	26	30	56
Immunology	23	21	44
Environmental Sciences & Engineering	16	27	43
Others	1183	1656	2839
Total	1740	2534	4274

An analysis of research areas (Table. 2) reveals that Pharmaceutical Sciences emerged as the top-most area for which this scheme was availed followed by Cancer/Radiation and Neurosciences etc.

One of the major mandates of the Council is capacity building of biomedical scientist of the country by providing them financial assistance for participating in International Conference/Training programmes/ Workshops etc. Out of total 4274 applications 1740 applicants were supported for International Conference/Training /Workshops etc during 2009-2013. Designation-wise analysis of the applicants who availed travel grants shows that the Research Fellows (JRF-SRF) had the highest share of 722 numbers. The other major section of researchers benefitted from the scheme was that of Faculty, which accounted for 489 of the availed grants (Figure 4).

In terms of gender mix, an analysis of the availed applications shows that 1122, out of total 1740 applications, or 26.3 % were males majority of whom were < 35 years of age. Similar trend was observed in the female segment with 618 availed applications and 14.5% being < 35 years of age.

Findings

An analysis of sample data during the period 2009-2013 shows that under the ICMR funded Financial Assistance to Non-ICMR Biomedical Scientists programme, out of the 1740 selected proposals:

- New Delhi led all other states in terms of the applications submitted (1320) and grants approved (361) followed by UP with 461 applications submitted and 89 approved.
- Amongst the institutes that led in the number of grants availed, AIIMS had a share of 15.1% followed by PGIMER, Chandigarh 16.3% and NIMHANS, Bangalore at 3.4% each.

- Pharmaceutical Sciences emerged as the top-most area for which this scheme was approved followed by Cancer, Neurosciences, Ophthalmology, Nuclear Medicine and Infectious Diseases.
- Designation-wise analysis of the applicants who availed travel grants shows that the Research Fellows (JRF-SRF) had the highest share of 722 out of 1854 applications followed by Faculty who accounted for 489 out of 1154, Senior & Junior Resident 203 out of 505, Md. Students 54 out of 143 of the approved grants.

References

Report of the Working Group on Health Research for XII Five Year Plan (September 2011). *Department of Health Research and Family Welfare*, 1-2.

Burroughs Wellcome Fund, *www.bwfund.org*

National Commission on Macroeconomic and Health

Health Information of India 2004, Government of India, Central Bureau of Health Intelligence, DGHS, MOH & FW, Nirman Bhavan, New Delhi-11 website:www.cbhidghs.nic.in

Growth of Literature in Biofuels Research: A Resource Analysis

Shantanu Ganguly*, P K Bhattacharya** and Tanvi Sharma***

*PhD, Fellow and Area Convenor, Knowledge Management Division, TERI

**PhD, Fellow, Knowledge Management Division, TERI

***Research Assistant, Knowledge Management Division, TERI

Abstract

The subject of biofuels has gained considerable importance in the past few years. It is today well known that biofuels use as a source of energy has many advantages; it is available easily and in abundance, helps in the reduction of GHG emissions, and has a positive effect on health as it is biodegradable and low in terms of its toxicity. Its use is also beneficial in terms of energy security and accrues benefits on economic terms. Application of biofuels in the transport sector has been proven with its efficient utilization as biodiesel. Biodiesel is an eco-friendly diesel and the past few years have seen it gain popularity within civil society.

This study is an attempt to understand current trends and development of literature in the field of biofuels in the last 10 years. This paper emphasizes upon research and development as well as technology breakthroughs in biofuels. This analysis has been carried out considering certain perspectives of biofuels. The growth of literature in each of the perspectives has been analysed after collating articles from different sources.

Methodology

This research involves gathering of relevant data from the specified journals and databases in order to analyse the development of biofuels in the past few years. This study aims to find out the most focused subject category under biofuels in the past 10 years as well as about the research and development (R&D) in biofuels. This analytical study has been carried out on the basis of articles published in leading internationally refereed journals and databases. Three major databases which were explored are Science Direct, JSTOR, and Springer. Amongst the journals perused are *Biomass and Bioenergy*, *Journal of Scientific and Industrial Research*, and various other journals with articles published on biofuels. Selection of journals was done not only keeping in view annual coverage on development of biofuels but was also based on the pattern followed in subject development within the databases and journals. The chosen articles consist of case studies, conceptual articles on of biofuels, and studies on biofuels with special reference to India. Articles by Indian authors as also articles about India have been given more weightage. These articles have been categorized under six main categories: Feedstock, Biofuel Processes, Development and Growth, Application and Use, Livelihood Aspects, and Environmental Issues

For this study, databases and journals have been scanned through to understand the development of biofuels in the past 10 years, from 2003 to 2012. The study attempts to solve three questions: (*i*) Which of the category under biofuels has gained maximum interest amongst authors? (*ii*) Which source out of the five Science Direct, JSTOR, Springer and two leading journals sources studied, has paid more importance to biofuels? and (*iii*) What has been the trend in development of biofuels subject in past 10 years?

Scope and Limitations

The limitation of this paper is that it is solely based on resources subscribed to by the Library and Information Centre, The Energy and Resources Institute (TERI).

A. Background

Energy is understood to be a key input to socio-economic development as it can play a critical role in enhancing productivity and reducing drudgery. This is particularly true for countries such as India, whose development levels and per capita energy consumption are low. Hence, it is important to holistically and objectively understand and assess the country's energy sector to identify its strengths and weaknesses so that the policies and interventions can be appropriately prioritized to further the country's development (Srinivas and Iyer, 2014).

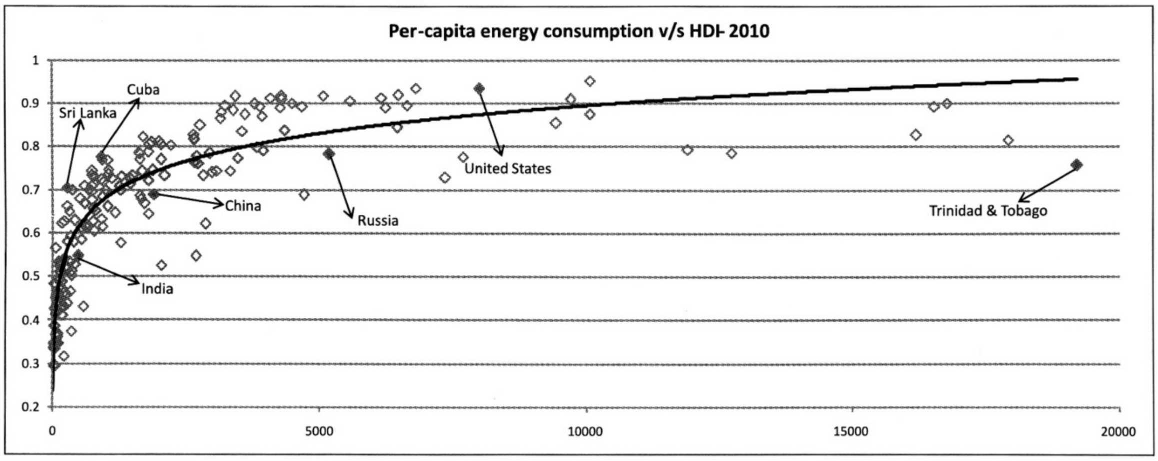


Figure: HDI levels and per capita primary energy consumption

Within a decade, R&D in the renewable energy sector in India has grown from being a fringe player to a mainstream actor of the energy sector. Installation of renewable energy for electricity has grown at an annual rate of 25 per cent, with data for January 2014 showing installed capacity at 30,000 MW. During this period, wind power installation has also grown 10 times and solar energy has grown from a bare minimal to 2,500 MW. Currently, renewable energy accounts for about 12 per cent of the total electricity generation capacity of the country and contributes to 6 per cent of the total electricity produced. Renewables therefore produce more than twice the amount of electricity produced by all nuclear power plants in the country. In 2012–13, electricity produced by renewables was equivalent to meeting the per capita annual electricity requirement of about 60 million people. Today, more than a million households in the country depend solely on solar energy for their basic electricity needs.

The growth of renewable energy has changed the energy business in India. It has, in many ways, democratised energy production and consumption in the country. Before the renewable sector became a significant player, the energy business was all about large fossil fuel based companies and grid-connected power; they dominate even today. At the same time however there is an alternate energy market in which thousands of small companies, NGOs, and social businesses are involved in selling renewable energy products and generating and distributing energy from renewable sources. This trend is likely to accelerate because of two key policies of the government.

The concept of biofuels is being developed and practised since 2003, but in 2004 growth and development of biofuels is minimal as there no article under this category published in any of the sources. Authors have been showing interest in writing on the developmental aspects of biofuels from 2006 onwards. Concepts in biofuels have also faced certain challenges that have been focused on by researchers and revealed in their papers.

Starting from 2006 going on till 2011, development of biofuels has come more into focus. There was a reduction in the articles from 2008 to 2009, which again took a high rise starting from six articles in 2009 going on to 13 articles in 2011. The sudden rise in the articles on development of biofuels can be credited to the National Biofuel Policy of India, which was approved by the Cabinet Committee in 2008 and released in 2009 (Raju et al., 2012). This shows that the government has started taking initiatives towards promotion of biofuels concept in the country. With this breakthrough, the authors were interested bringing out more studies on development as well as the growth of biofuels.

B. TERI Library and Information Centre: A Brief Knowledge Sketch

TERI was established in 1974 as an institution committed to deal with every aspect of sustainable development. All activities in TERI move from formulating local- and national-level strategies to developing global solutions for critical energy and environment issues. While in the initial period the focus was mainly on documentation and information dissemination activities, research activities in the fields of energy, environment, and sustainable development were initiated towards the end of 1982. Over the last 30 years, TERI has created an environment that is enabling, dynamic, and inspiring for the development of solutions to local and global problems in the fields of energy, environment, and current patterns of development, which are largely unsustainable. The global presence and reach attained by TERI are not only substantiated by its presence in different parts of the world but also in terms of the wide geographical relevance of its activities.

As TERI grew in size and expanded activities, it was considered essential to maintain and nurture the intellectual capital that it has assimilated over the years due to three main considerations, viz. (i) improving efficiency; (ii) avoiding knowledge loss; and (iii) stimulating knowledge growth and creation. All these were possible because of the robust Knowledge Management system at TERI.

B.1 ENVIS Centre on Renewable Energy and Environment

The Environmental Information System (ENVIS) network was established as a planned programme under the Ministry of Environment and Forests (MoEF), Government of India, in December 1982. Initially, MoEF identified 25 organizations to host ENVIS Centres. Thereafter, since January 2002, ENVIS started implementing a series of World Bank-assisted Environment Management Capacity Building Technical Assistance Project (EMCBTAP). Over the years, the MoEF has identified a few organizations as centres of excellence who host subject-specific ENVIS centres to disseminate environmental information. In addition to this, ENVIS centres are also hosted at State Environment and Forests Departments for wider dissemination of state environmental information. Presently, there are a total of 67 ENVIS centres spread across the country which include subject-specific and state-level centres.

Primary objective of the project is to support R&D and create awareness, collect information, and disseminate publications on environmental issues across India among the research community, students, and policy-makers. Further to this, the project also aims to develop IT expertise among the ENVIS Centre(s) professionals.

TERI has been hosting the ENVIS Centre on Renewable Energy and Environment since July 1984. The Centre is presently located within the TERI Library and Information Centre, Knowledge Management Division, as a dedicated centre for environmental information collection, collation, and dissemination. The major objectives of the Centre are:

- Collection and dissemination of information to support and promote research and development and innovation among researcher, policy makers, academicians, and other stakeholders
- Identify and bridge data gaps
- Build and maintain databases
- Bringing out publications and disseminate to stakeholders

The centre successfully identified data gaps in the areas of environmental impact of power, sustainable transportation, fossil fuels and environment, hazardous waste management, pollution control technologies, hazardous waste management, environmental laws and regulations, environmental economics, etc. Conscious efforts are being made to bridge these gaps by building online databases, bringing out publications and reports, organizing seminars, and query response to users. Besides, the centre also provides journal contents services, bibliographic services, document delivery services, and other related activities to meet its objectives.

Figure 1: ENVIS website

The dynamic ENVIS website today (www.terienvis.nic.in) (Figure 1) hosts updated resources and databases that it generates and collects for users and maintains periodically. The ENVIS website has several sections which highlight centre's activities which include recent news, government regulations, updated links, and glossary of terms, technologies, case studies, statistics, publications, bibliographies and databases. The centre also conducts user-interactive annual workshops for popularizing the ENVIS centre to understand their needs. This paper is a part of the complete report available on the ENVIS website.

C. Research Methodology

This research involves gathering of relevant data from the specified journals and databases in order to analyse the subject development of biofuels in the past few years. This study is about finding the most focused subject category under biofuels in the past 10 years as well as about the development of biofuels research. Research has been carried out on the basis of the articled collected. The three major databases accessed are Science Direct, JSTOR, and Springer, while journals referred to are *Biomass and Bioenergy*, *Journal of Scientific and Industrial Research*, and various other journals that had a collection of articles on biofuels. An analytical study has been done after a selection of these journals and databases which explored the pattern followed in subject development within the databases and journals.

C.1 Resource Analysis

The up-growth of the biofuels subject has been studied using different information sources. In this report, three major sources of information have been analysed, i.e., books, journals, and databases. Books have been analysed from the collection of available books on amazon.in as well as in TERI library and Information Centre. The journals and databases have been studied for articles that have contributed to the research dimension of biofuels.

C.1. (a) AMAZON.IN

In the Amazon.in book database, a search was made for "Energy" and "Biofuels" related books. A total 468 books were identified out of which 284 books have been published on "Energy" and "India" in the past 10 years. Several aspects of energy have been addressed by both Indian as well as international authors. The coverage of biofuels was minimal as shown in Figure 2.

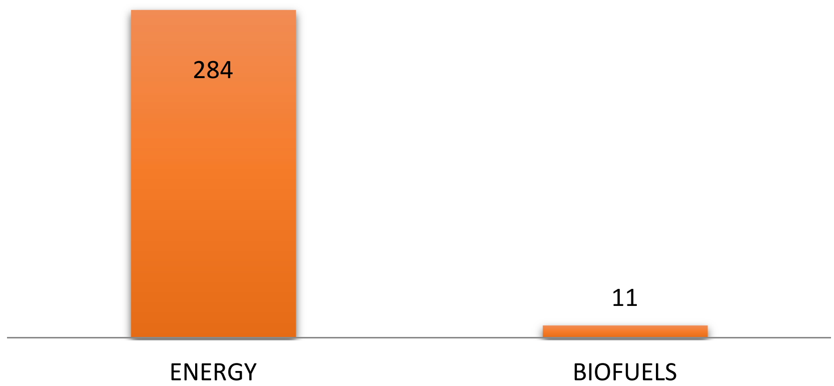


Figure 2: Books on Amazon.in on Biofuels and Energy

C.1. (b) Scientific and Scholarly Publications

Analysis was based on select journals that are subscribed by TERI. Some of the select journals referred to are *Renewable Energy*, *Energy Sources*, *Journal of Scientific and Industrial Research*, and *Biomass and Bioenergy*. Out of these, the *Journal of Scientific and Industrial Research* had published 61 articles and *Biomass and Bioenergy* had published 44 articles on Biofuels, whereas. While *Energy Sources* had published a large number of articles on Biofuels but it had not published any article by an Indian author. In *Renewable Energy*, there was not much on the subject. An analysis of published papers on Biofuels has been depicted in Figures 3.

i. Journal of Scientific and Industrial Research (JSIR)

Journal of Scientific and Industrial Research comprises articles dealing with different aspects of science and technology, including industry. This journal includes articles on various issues of industrial development, industrial research, technology management, technology forecasting and others.

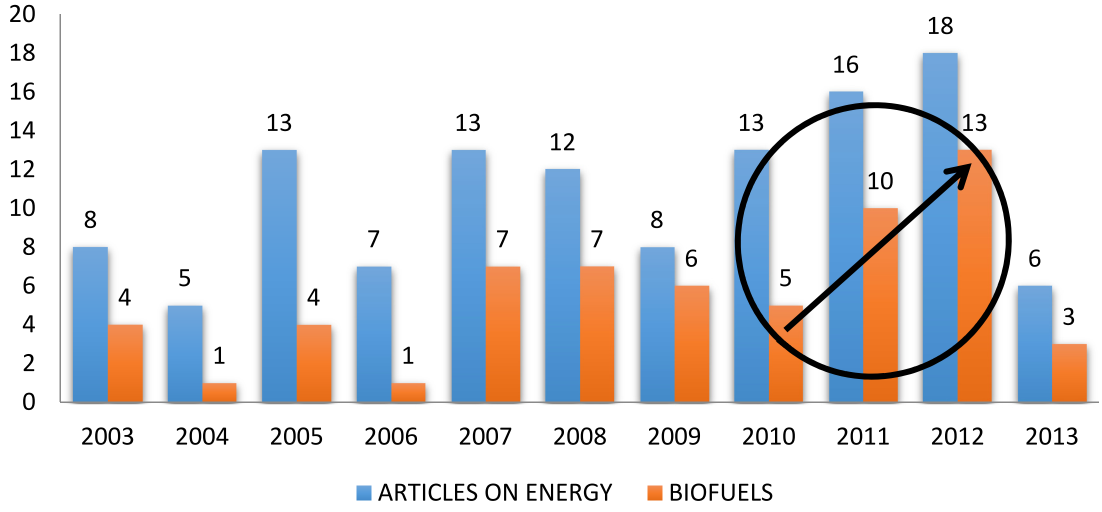


Figure 3: Articles published in Journal of Scientific and Industrial Research

From 2010, studies on biofuels have taken a steep rise from 5 articles to 13 articles in 2012 (Figure 3).

ii. Biomass and Bioenergy

Biomass and Bioenergy is an international journal by Elsevier, which collates worldwide studies on biomass as a renewable source of energy.

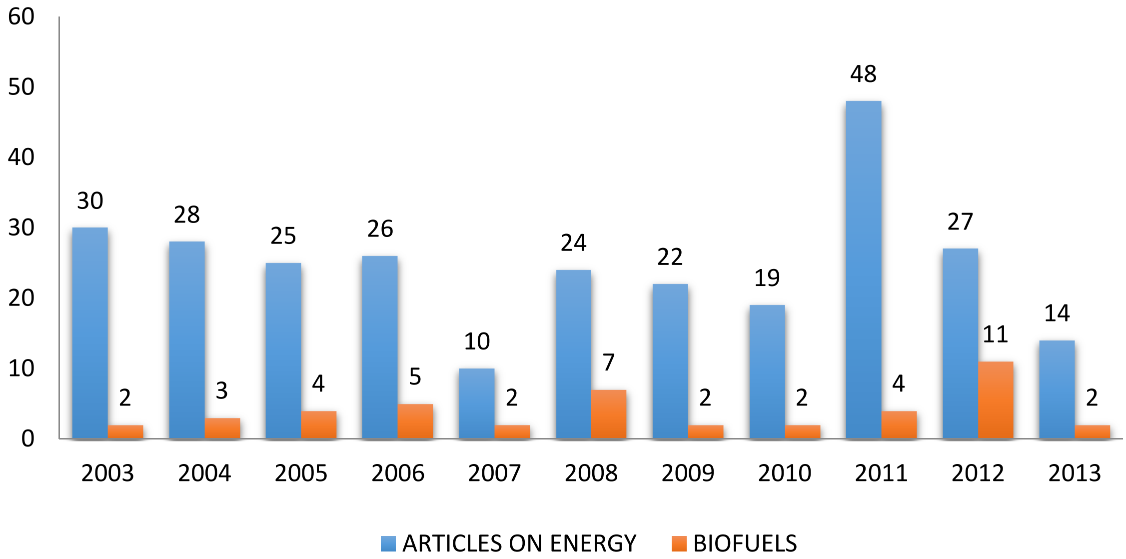


Figure 4: Biomass and Bioenergy

In the context of biofuels as well as biofuels in India, this journal has published 273 articles on energy in past 10 years. Publishing 44 articles specifically on biofuels (Figure 4), Elsevier has published most of the articles either on feedstock or on biofuel processes. Out of all the articles on biofuels, 19 articles are on feedstock whereas 13 articles fall in the category of biofuels processes, i.e., a total of 73% of the articles in *Biomass and Bioenergy* belong to the category of feedstock and biofuel processes.

iii. Comparative Analysis of JSIR and Biomass and Bioenergy

The *Journal of Scientific and Industrial Research* (JSIR) has published 34% of its articles on application and use of biofuels, but on the other side, it has also given importance to studies on feedstock and biofuel processes with 24% and 21 % articles respectively as shown in Figure 5.

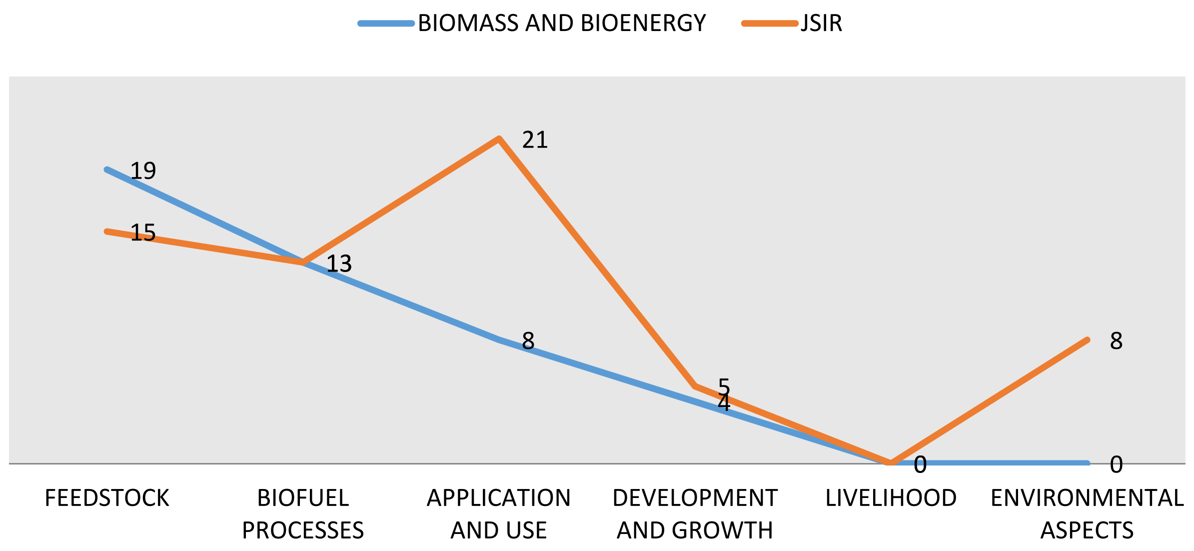


Figure 5: Comparative between Biomass and Bioenergy and JSIR

Biomass and Bioenergy has published 25% of biofuel articles in 2012 with a stockpile of 11 articles out of 44. Though it has dealt with the major concepts of biofuels on a worldwide base, in context to India, its contribution has been precisely limited. As can be seen from the graph, (Fig 4) since 2003 till 2010, the share of articles on biofuels has been very low. 2012 has been a significant year for biomass and bioenergy where Indian authors have greatly contributed to the study of biofuels. There has been an upsurge in biofuel articles in this journal from 2011 to 2012.

C.1. (c) Databases

The up-growth analyses of biofuels articles using databases is done from the collection of three most popular databases, i.e. Science Direct, JSTOR and Springerlink. Science Direct, JSTOR and Springerlink are most popular e- databases that have a huge collection of articles from different books and journals. They comprise of articles on different subject areas such as energy, economics, social sciences and many more. In the past ten years, these three databases altogether have managed to have a stockpile of 4848 articles on energy as a subject area and 202 articles on biofuels. The Fig 6, Fig 7, Fig 8 shows the upgrowth of articles in the biofuels sector in comparison to Energy sector as a whole.

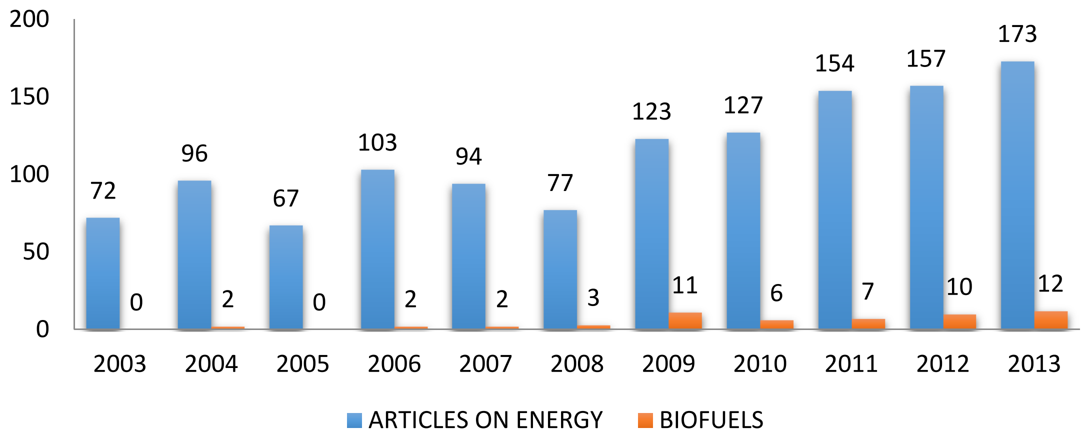


Figure 6: Articles published in *Science Direct* database

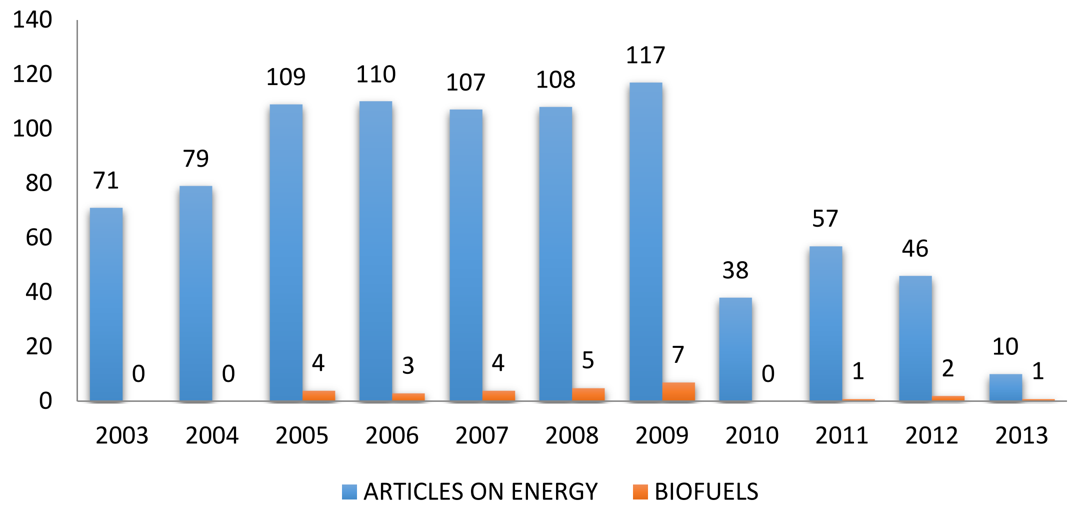


Figure 7: Articles published in JSTOR database

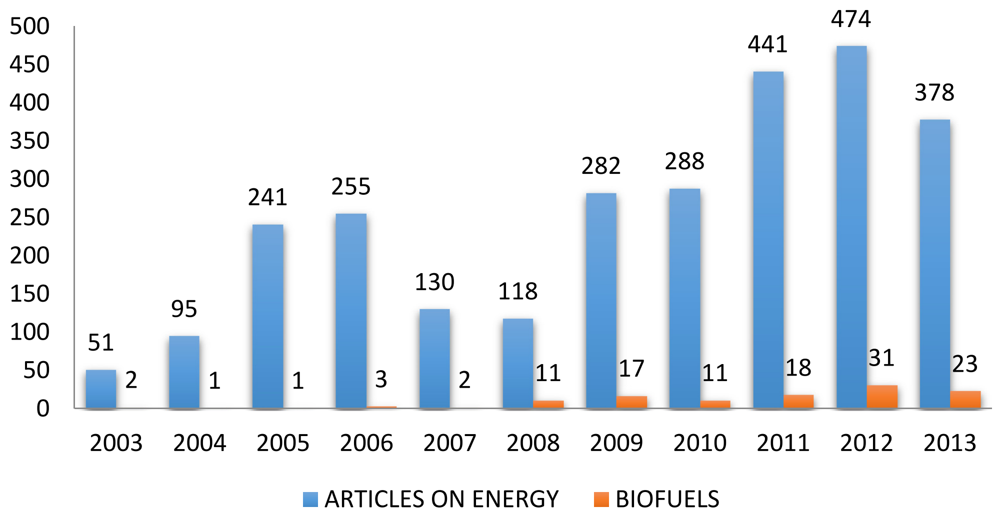


Figure 8: Articles published in Springerlink database

With a collection of 1243 energy articles with 229 articles on renewable energy, Science Direct has a collection of 55 articles that are thoroughly based on biofuel subject (Fig 6). With a collation of 120 biofuel articles out of 700 renewable energy articles, SpringerLink database has a collection of 2753 (Fig 8) energy articles that counts to be almost twice of the articles in Science Direct. JSTOR has brought out very few articles on biofuels subject, with 27 biofuels articles and 852 energy articles in the past ten years (Fig 7). Out of 55 articles 83 % of the articles in science direct are found to have been collated after 2009, with 2009 and 2013 giving the highest share of 11 and 12 articles respectively. Since 2010, there has been a continuous rise in the articles on energy as well as on biofuels (Fig 6). Similar scenario is formed in springer. Springer had a collection of 120 articles on biofuels in past ten years, out of which nearly 93% of the articles were found to be between 2008 and 2013. With 2012 having 31 articles, Springer brought out most of the articles on energy also in 2012 itself (Fig 8). The focus of Science Direct and Springer shifted towards biofuels 2009 onwards, when biofuels concept became prominent and various authors got interested in writing on biofuels as a subject. Various aspects of biofuels gained importance, in turn giving rise to the articles in various journals and hence in different databases. Though JSTOR has not shown much interest in biofuels after 2009, in 2009 itself, it had most of the articles on biofuels amongst past ten years. Out of 27 articles, 7 articles were found to be in 2009 (Fig 7). Amongst all three databases, year 2009 has been a year when all the three have shown considerable movement in the subject of biofuel.

D. Biofuel Subject Development: A Research Trend

The concept of biofuels has gained considerable importance in the last one decade. India has been focusing on bringing out the most environment friendly fuel that can replace fossil fuels majorly in transport sector. Also to limit its imports, the country has been focusing on the biofuels alternative that can readily be made available in India. The government of India has also shown keen interest in development and promotion of these biofuels through bringing out policy measures and the research and development in this perspective. This has interested the authors to write upon various aspects of biofuels with regard to India and also about the whole concept of biofuels that leads to a further research on the subject. A total of 460 articles published till 2012 and additional 29 articles in 2013 have been collected under the biofuels subject. There has been a consistent increase in the number of articles from 2003 to 2012 (Fig 9).

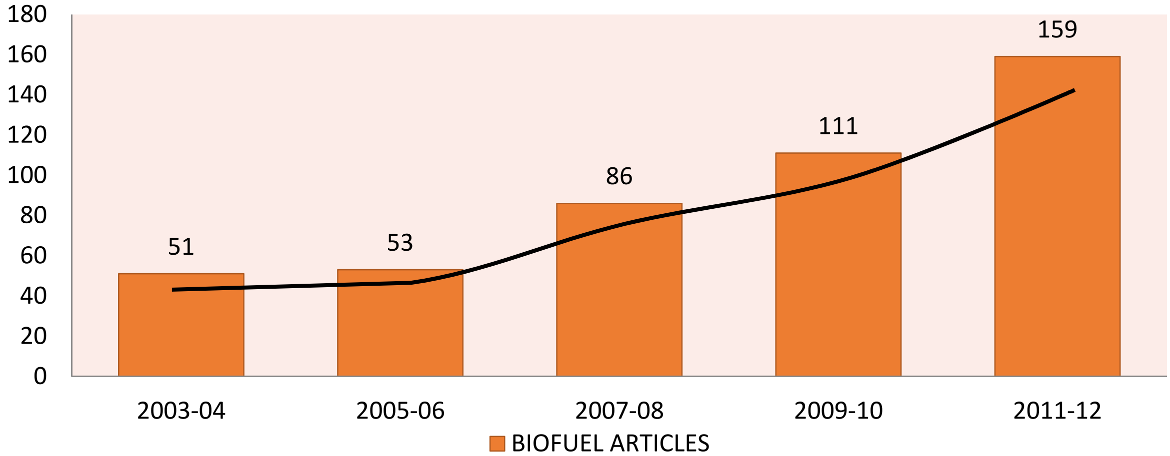


Figure 9: Biofuels Subject Development

Articles for the study have been collected for the past ten years. These articles consist of case studies, articles on the concepts of biofuels as well as studies on biofuels in context to India.

Articles by Indian authors as well as articles about India have been given more focus besides other articles. As mention in the Table No. 1 below, accessed articles were categorized broadly into six categories.

Table 1: Different categories and articles accessed

Category	No. of Articles
Biofuel Processes	128
Feedstock	127
Application and use	93
Development and Growth	77
Environmental Aspects	48
Livelihood Aspects	16

D.1 Feedstock

Collection of articles under feedstock section is based on the content of articles that inherit information about the feedstock options that are available for the production of biofuels as well as on plantations of feedstock required for biofuels. Studies dealing with different aspects of feedstock such as the plants and biomass that can be used as a raw material for biofuels have also been considered under feedstock category.

In this decade, most of the studies have been on biofuel processes followed by feedstock. It has been analyzed that in 2012 with 22 out of 117 articles being on some or the other aspect of feedstock, the five sources studied in this research, have brought out maximum of its articles on feedstock in 2012 (Fig 10). Followed by 2012 is 2011 which had a share of 19 articles being thoroughly on feedstock. The statistics show that most of the popularity on feedstock was gained in the period of 2009 to 2012. Starting from 2003 having a downfall, the studies on feedstock took a boost in 2009 continuing till 2012. The focus was centered towards bringing out the suitable feedstock for biofuel production.

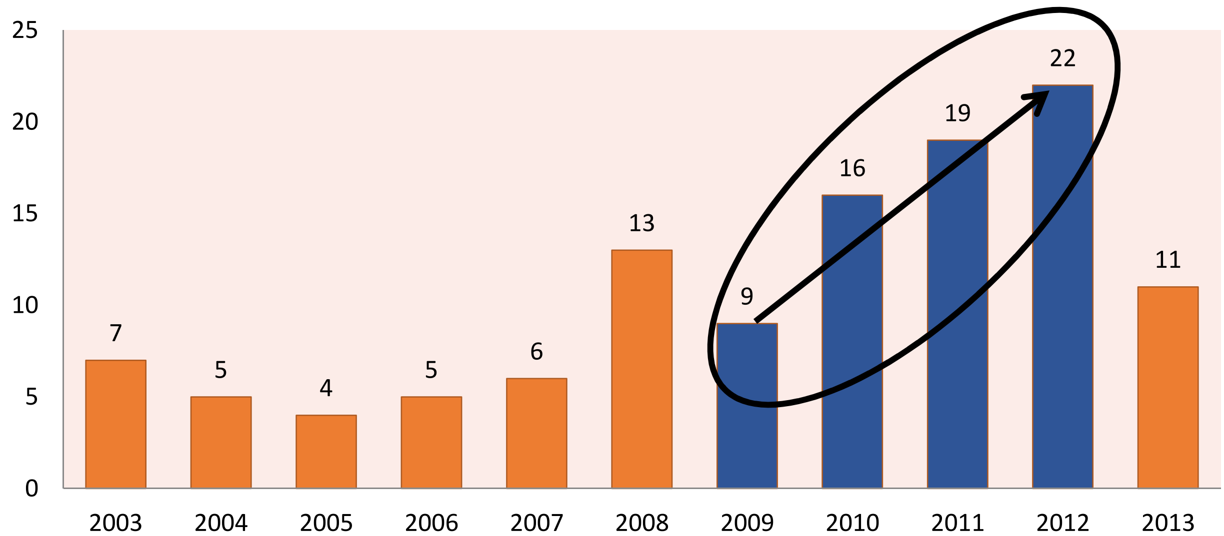


Figure 10: Publication of articles on feedstock

One of the major sources for biodiesel in India which was identified by planning commission was *Jatropha curcas*. Planning Commission of India identified *Jatropha* which is a non-edible oil bearing tree capable of producing oil that is easily convertible in to biodiesel (Jain S.K., et al, 2011). The identification of such a tree borne oil seed contributed to studies on feedstock.

D.2 Biofuel processes

It comprises of the studies done on processes involved in production of biofuels. Articles on different experiments in production of biofuels through different processes are also a part of this category. Besides this, studies on technologies associated with the concept of biofuels are also categorized under this section. Wherever authors have written about biofuels in context to processes and technologies, it has been classified under 'biofuel processes'.

Biofuel Processes has got considerable attention in past ten years. Authors have written about technological aspects of extracting biofuels every year. Some or the other process of extracting and blending of biofuels such as fermentations, thermochemical conversions and so on have been given attention each year. Year 2012 has got exceptionally high statistics in the studies on biofuel processes.

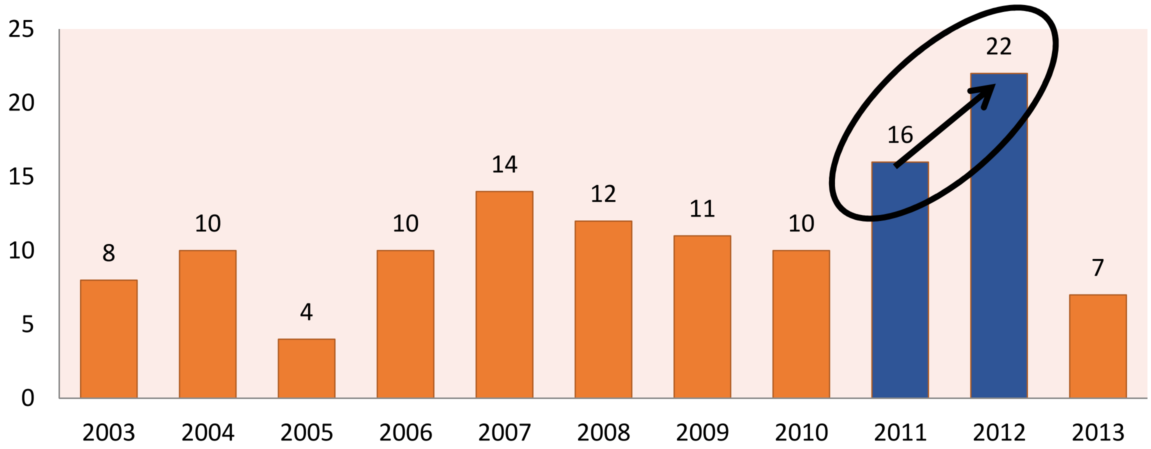


Figure 11: Articles published on biofuel processes

Identification of *Jatropha* as a biodiesel crop also brought out an opportunity to analyze the different technological methods to extract bio oil from its seeds. Studies on the methods of extracting bio-oil got a push through the popularization of the tree-borne oil seed *Jatropha*. Out of 124 articles on biofuel processes in past ten years, 2012 has got 22 articles which counts for nearly 18% of the total articles. Most of the studies on blending of biofuels and technologies for extraction of bio-oils have come up in the phase of 2011 and 2012 having 31% of the total share.

D.3 Development and growth

Articles on policy issues, and awareness & understanding of biofuels have been classified under development and growth of biofuels. Besides this, studies on the challenges posed in front of biofuels conceptualization have also been categorized under development and growth of biofuels, as they also form a part of growth process of biofuels.

The trend shows that the concept of biofuels has been getting developed and practiced since 2003. It has been into light in the past one decade. Whereas, year 2004 is an exception as there is no article under this category. Authors have been showing keen interest in writing on the developmental aspects of biofuels. There was a decline in in 2008 and 2009, but slowly gained importance in 2011. This increment is due to the formulation of National Biofuel Policy of

India, which is being approved by the Cabinet Committee in 2008 and the report was released in 2009 (Raju S S., et. Al, 2012). This shows that government was also interested in developing the biofuels concept in India which interested the authors to bring out more studies on development as well as the growth of biofuels.

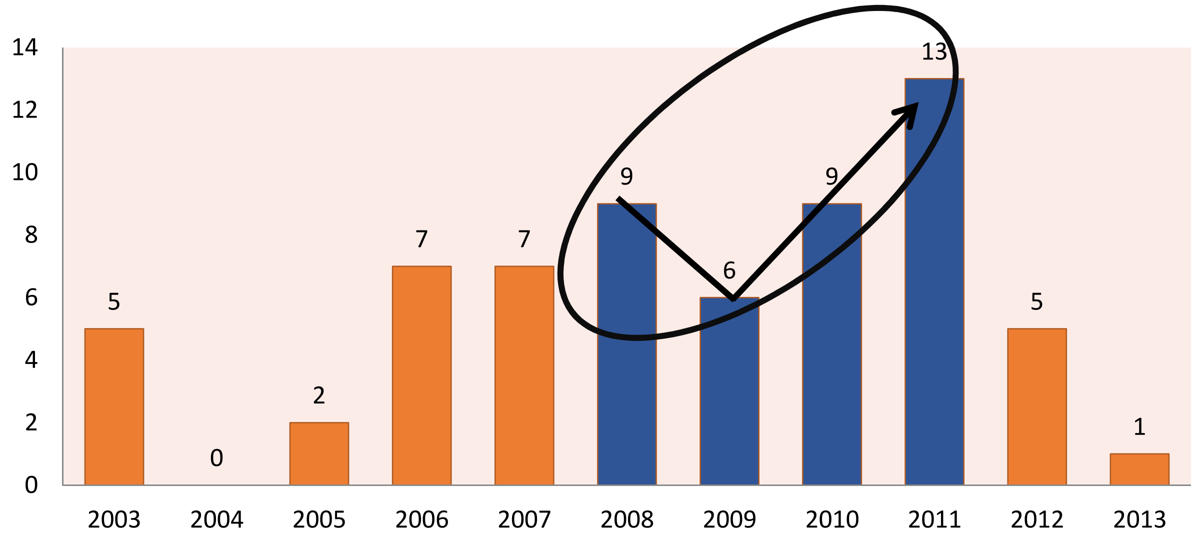


Figure 12: Articles published on biofuels development and growth

D.4 Application and usage

This involves articles about the testing of biofuels in different situations, as well as articles on actual practicing of biofuels. The studies on outcomes and consequences of application of biofuels have also been classified under application and usage. The development of biofuels is connected to the application of biofuels such that more the development of biofuels, more the application and the usage. From 2009 onwards more studies on application of biofuels has been observed, (Fig 13).

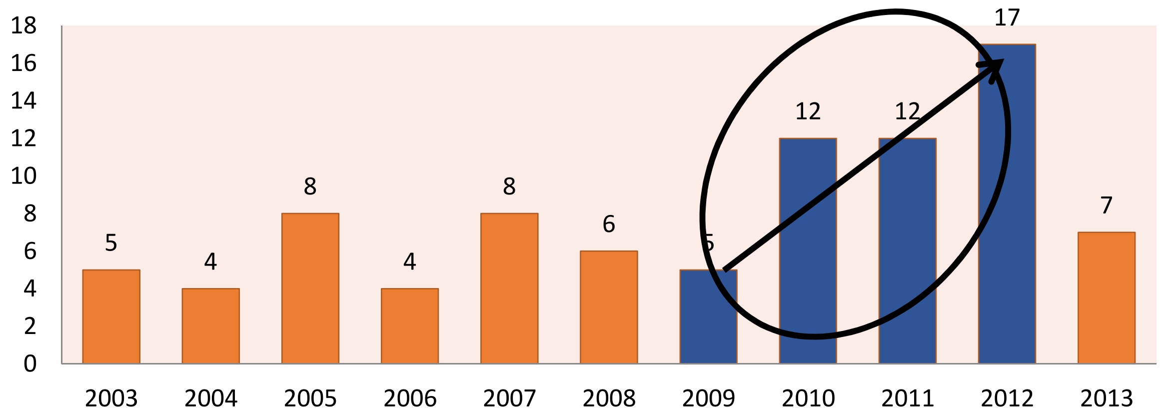


Figure 13: Articles published on application and use of biofuels

Authors wrote much about the trials as well as the actual applications of the biofuels. The national biofuel policy of India itself envisions biofuels as the most likely substitute to the diesel demands which will not only glean the environmental benefits but will also contribute to the rural development with its wide scale application and use (Raju S S., et. al, 2012). Starting from 2009 reaching to 2012 there has been a threefold increase in the articles on biofuel application.

D.5 Livelihood

The articles that address the sustenance factors have been categorized under the livelihood sector of biofuels. Biofuel papers that have focused on issues related to poverty as well as employment of population have been kept under this category of articles.

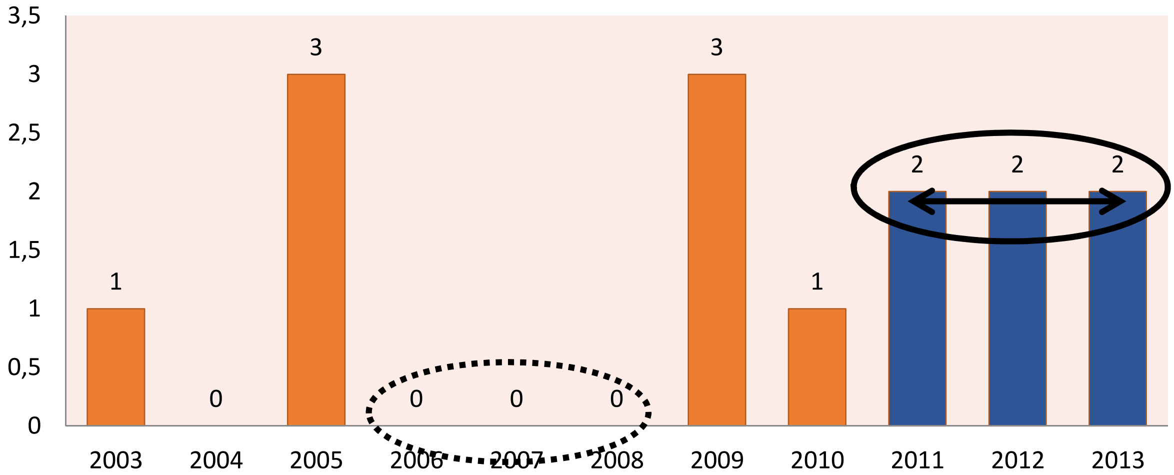


Figure 14: Articles published on biofuels as livelihood

Pattern of livelihood studies under biofuels is very distinct. Articles on livelihood concepts have gone up exceptionally in 2005 as well as in 2009. Although the total articles on livelihood aspects stand to be just 14 in number, it shows that the *'livelihood'* perspective has not been much in light. It has got attention but not as much as the other perspectives under biofuels. Fig. 14 depicts that between 2006 and 2008 no articles published on livelihood. Whereas in 2011 and 2012, some of the lead aspects like poverty issues, employment opportunities and other aspects have been gained importance. A total of 6 articles out of 14 have been brought out in the period of 2010 to 2012. The popularization of *Jatropha* as a biodiesel feedstock has been in focus since 2003, but in the recent years, it has interested the authors to write upon the pros and cons of *Jatropha*, such as the *employment opportunities*, addressing the *poverty issues* have gained attention of various authors on livelihood aspects.

D.6 Environmental Issues

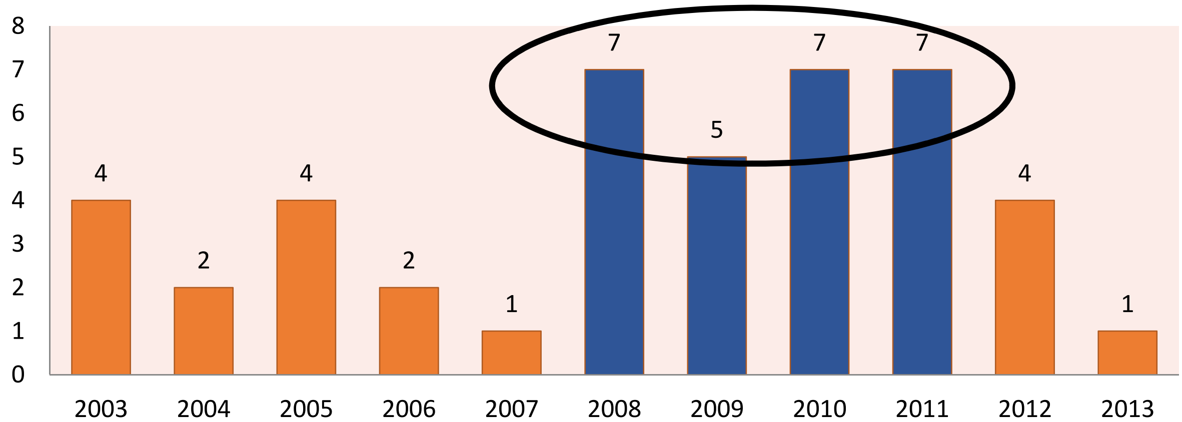


Figure 15: Articles published on impact of biofuels on environment

Environmental issues associated with the concept of biofuels have gained maximum attention in the years 2008, 2010 and 2011. There are 50% (i.e. 21 out of 44) have been published in these three years.

D.7 Comparative of literatures accessed and growth of biofuels research trend

Studies on different variables of biofuels have been very plangent in the past ten years. Amongst various aspects of biofuels, namely feedstock, application and use, development and growth, biofuel processes, environmental issues and livelihood, maximum of the studies have been on feedstock as well as biofuel processes. With a total of 237 articles out of 460, nearly 52% of the articles belong to these two categories. However, with a share of 162 articles, development and growth & application and use have got 35 % of the studies in their favor. Articles on environmental issues and livelihoods have been very low in numbers and have shown a static appearance. There have not been many articles in these two areas. Instead of this, period of 2009 to 2012 has been the most vigilant period in context to biofuels. In this period, each aspect of biofuels, as listed above, has got considerable importance from authors. Various authors have written about biofuels, as a subject as well as about biofuels in India.

Table 2: Components of Biofuels and articles accessed

	Feedstock	Application and Use	Development and Growth	Biofuel Processes	Environmental Issues	Livelihood
2003-04	12	9	5	18	6	1
2005-06	9	12	9	14	6	3
2007-08	19	14	17	26	8	2
2009-10	30	19	23	21	13	5
2011-12	46	32	22	42	13	4

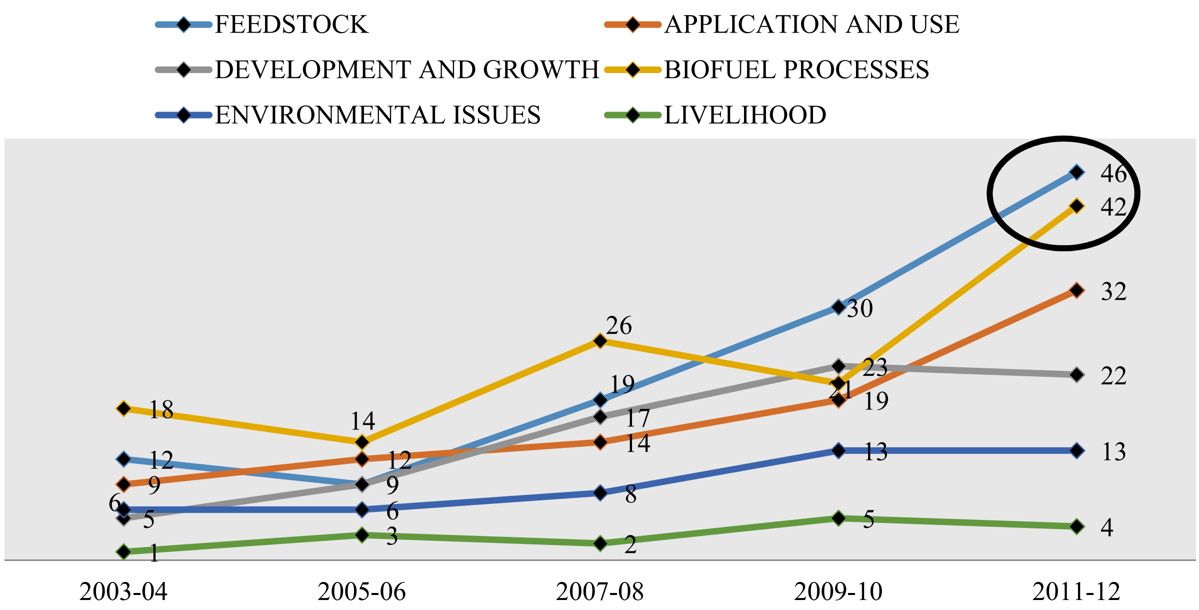


Figure 16: Different components of Biofuels

E. Conclusion

This study collectively provides a clear picture of the scenario that the development of biofuels subject has followed in the past ten years. Starting from 2003 reaching 2012, the development of biofuels as a subject has been studied, through the available databases and journals. The analytical study for past ten years, tries to solve the following questions:

1) Which of the category under biofuels has gained maximum interest of the authors?
2) Which source out of the five sources studied, has paid more importance to biofuels? and
3) What has been the trend in development of biofuels subject in past ten years?

From the study it has been analyzed that year 2011 and 2012 have been dominant years in the context of biofuels. Authors have shown keen interest in writing about biofuels in these years. Most of the studies on biofuels as an aspect and particularly on biofuels in India have come up in 2012 making it a milestone for biofuels concept particularly for India. Besides this fact, this period has been eminent for journals and databases that deal with the articles on biofuels. Journal of Scientific and Industrial research (JSIR) (fig 3) starting from a collection of 4 articles in 2003 reached its maximum level in 2012 with 13 articles on biofuels. From 2003 to 2012 the stockpile of articles in JSIR has shown a wavering appearance. Compared to this journal, if we look at the Springer database (fig 8) the steep rise from 2010 to 2012 depicts that there has been a sudden rise in the collection of articles on biofuels in this database. Having 11 articles on biofuels in 2010, the number rose to 31 articles in 2012. There has been a thrice increase in the collection of studies on biofuels in Springer database. Besides having a dip in the collection of articles in 2010, the five studied database and journals have gained a high rise in year 2012.

The analyses state that most of the studies have come up to be in Feedstock and biofuel processes. These two aspects of biofuels have been paid much attention. Amongst the information sources studied, Springerlink database and the journal of scientific and industrial research have had the maximum studies about biofuels. Though JSIR has shown more number of articles, biomass and bioenergy has also stood significant in accumulation of articles on biofuels in relation to India. With time, the concept as well as the application of biofuels has emerged as a pertinent approach to the question of depletion of fossil fuels and environment sustainability. In context to India, the national biofuel policy of India has also played a major role in giving a thrust to the research dimension of biofuel subject. The continuous improvements, advancements and immense deliberations in the concept of biofuels prove the authenticity of biofuels as an alternative source of energy.

References

Srinivas A. and R. K. Iyer, 2014. *A Comprehensive, Multi-dimensional Energy Index*. Pune: Prayas, p. 3.

Raju S. S., S. Parappurathu, R. Chand, P. K. Joshi, P. Kumar, S. Msangi, 2012. *Biofuels in India: Potential, Policy and Emerging Paradigms*, New Delhi: NCAP Publication Committee, Chapter 3: ‘Policy Framework and Enabling Factors’, pp. 17–32. Available at http://www.ncap.res.in/upload_files/policy_paper/pp27.pdf (last accessed on September 5, 2013).

Jain S. K., S. Kumar, and A. Chaube, 2011. “Jatropha Biodiesel: Key to Attainment of Sustainable Rural Bioenergy Regime in India”. Available at http://scholarsresearchlibrary.com/aasr-vol3-iss1/AASR-2011-3-1-425-435.pdf (last accessed on September 18, 2013).

Comparative study of outputs and scientific cooperation of world's countries in Biomedical engineering field in Science Citation Index in the years 2002-2011 with an emphasis on co-authorship networks

Soheila Bagheri* and Mohaddeseh Dokhtesmati**

*M.A in information and knowledge science, Islamic Azad University, Tehran science and research branch, Tehran, Iran
shlbagheri3@gmail.com

**M.A in information and knowledge science, Islamic Azad University, Tonekabon branch, Tonekabon, Iran
mdokhtesmati@gmail.com

Abstract

Due to the importance of Biomedical engineering and its close relationship with the society's health and determining the world's countries' status in this field of science, the present study has compared the scientific outputs of world's countries and their scientific cooperation in the Biomedical engineering field at the science Citation Index (SCI) in the years between 2002-2012. The research methodology is descriptive-survey and Scientometrics indicators have been applied in it. The research sample consists of 12044 academic and research outputs in the field of Biomedical engineering which have been taken from *web of science* website in the years 2002-2012. Also, in order to illustrate the co-authorship network of 20 superior countries in Biomedical engineering productions, the NodeXL software was applied.

The research findings showed that among all continents, Europe has the most and Africa has the least contribution in scientific productions in Biomedical engineering. United States has got the most scientific outputs and the most cooperation share in scientific productions in Biomedical engineering. From Asia, China is the pioneer in these cases. The best research center in this field (Calif System University) and the best author (Kaplan DL) are also from America. The most topical areas were related to materials science, Biophysics and sport sciences in which more than 4 authors have cooperated in most documents. Other findings showed that the number of scientific productions in Biomedical engineering has ascending growth up to the year 2010 and in the year 2011 has a 6% fall. Most of the documents are articles and the least type of documents to book reviews.

Results regarding the co-authorship networks show that America, Germany, England, Japan and France were chosen as the first options of scientific cooperation by other countries. In their scientific productions, America, Canada, Germany, France and Japan made use of other countries' help more than the rest. Canada, America, Germany, France and Japan have the highest status in the network and their non-existence led to a break in relationships between other countries. Taiwan, Israel, Brazil, India, South Korea, Portugal, Spain and Singapore have been more influential in distribution of network information. More than other countries, Canada, Germany, America, England, Japan and France have been connected with active countries which have higher grades and more significant social role, respectively. Also, Brazil, Israel, Taiwan, Portugal, Spain and Singapore have higher clustering coefficient, respectively.

Introduction

One of the most dimensions of permanent development in every country is producing the scientific information. Information means power and those countries that are developed in scientific productions will be considered as powerful. Each country's scientists' cooperation in improving science at international level is one of the important topics in the discussion of science production (Osare, Norouzi and Keshvari, 2010).

Being aware of scientific and research outputs, scientific productions and developments undertaken in this area presents a thorough picture of scientific activities of researchers and authors in this area and makes it possible to identify the weak points and strong points of researches done. Nowadays, one of the most valid methods for evaluating the scientific productions and research outputs is using of Scientometrics indicators which are briefly mentioned as the knowledge of measuring science (Bookestain, 1994). Since the science development is resulted from cumulative activities (Nikzad, 2010) and mixed experiences and talents of members of a group can lead to the production of a higher qualified article, compared to an author who does everything by himself (HasanZade&Baqaee, 2009; He, 2009) and also leads to new scientific thought, decrease in research costs (He, 2009) , performance improvement and more research efficacy (Cheong &Corbitt, 2009) and can multiply the speed of different countries' development (Velayati, 2008), so the study of quality and quantity of cooperation among scientists is a topic which have been noticed by Scientometric researchers for several decades (Nikzad, 2010). So that in this regard, the research findings of Teodorescu & Andrei (2011), Ardanuy (2011), and Ordonez & Cozzens &Garcia (2010) showed that multi-author articles are more relied on and those authors who had more scientific cooperation with other researchers produced more articles.

One of the science branches that nowadays have come to the focus of researchers is the Medical Engineering major. This major in the two recent decades in the industry sector and among its different areas has had the highest growth, especially compared to classic industrial areas. While almost all the classic industrial areas have an annual growth lower than 5%, Medical Engineering and a few numbers of other technological areas have an annual growth up to 10%. In addition, extra importance of Medical Engineering is due to its direct relationship with medical major (scientific pole of medical engineering, 2007). This major is the applying of technical and engineering sciences to help doctors to diagnose and treat the diseases. So that it expands the accuracy and diversity in diagnosis. Also, without some special machines, recognizing some diseases is impossible (Mashhad Medical engineering, 2012). It can be said that the growth of medical science is owed to the existence of laboratory machines and equipment (KhosroAbadi and ZamaniNejad, 2007).

Regarding the role of medical engineering in providing more useful and qualified services in medical area, there is no doubt that considering the scientific cooperation of researchers in this area for writing and publishing scientific documents and its quality and quantity can be very precious. Therefore, the present study seeks to investigate the scientific and research outputs and the amount of scientific cooperation of international science producers in Medical engineering field in Science Citation Index at ISI website and to illustrate and examine their co-authorship networks.

Research methodology

The research methodology is analytic-descriptive-survey and Scientometrics indicators in analyzing multi-authorship networks have been applied in it. The research sample consists of 12044 academic and research outputs in the field of Biomedical engineering which have been taken from *web of science* website indexed in the years 2002-2012. To extract the data, version 5/7 of *web of science* website has been used.The method of data gathering in the present research was so that first, through the following tag:

> TS=biomedical* OR TS=bioinformatic* OR TS=biomechanic* OR TS=bioelec* OR TS=bioengineering* OR TS=biomaterial*

In Advanced search in Science Citation Index and by limiting the time distance to the years 2002-2011, 94645 records were obtained which by choosing the topical group of Biomedical engineering from *Web of Science Category* section, this number was reduced to 12309 records. Since the amount of scientific cooperation for books is less than that of articles (Moody, 2004) and scientific articles are more suitable for analyzing the cooperation, by limiting the results to different kinds of articles, conference papers, conference abstracts and editions, the number of records was reduced to 12044 scientific documents. The documents' bibliographical information from the *Marked list* part was saved in (win) Tab Delimited format for counting and analyzing in Excel and NodeLX software. For counting the data (frequency of medical engineering documents in each continent and world's countries, their frequency based on the type of the source for each year, the status of medical engineering's highly- worked topics based on multi-authorship, peer countries of the world's superior countries in scientific productions of medical engineering, the world's best institutions in this field, investigating the group cooperation status of the world's ten best authors in this field along with the total amount of the received references and their nationality) the Excel program was used. Furthermore, for illustration and network analysis, NodeXL software was applied. Also, for determining the number ofauthors in each record, with regard to the high amount of data and for speeding the research, the following formula in the Excel was applied:

=LEN(B1) - LEN(SUBSTITUTE(LOWER(B1); ";"; ""))+1f_x

It should be noted that due to the high amount of information, only ten first ranks of each case are mentioned in the above tables.

Data analysis

Table 1. Frequency of five continent's biomedical engineering documents

Continent	*Number of countries*	*Records*	*Percentage of records*
Europe	40	5661	47
U.S.A.	13	5568	46.23
Asia	28	3076	25.53
Oceania	2	466	3.86
Africa	7	44	0.36

The results of Table 1 show that Europe with 47% (5661 documents) has the most contribution in scientific productions in Biomedical engineering. Then, the most contribution belongs to America (46.23%, 5568 documents). Africa with less than 1 percent (0.36%, 44 documents) has the least contribution.

Table 2. Frequency of world's Top ten countries in biomedical engineering documents

Country	*Records*	*Percentage of records*
U.S.A.	4427	36.75
China	990	8.21
Germany	869	7.21
Canada	818	6.79
England	754	6.26
Italy	697	5.78
France	615	5.10
Japan	588	4.88
Netherlands	399	3.31
Australia	393	3.26

According to Table 2, the five superior world's countries in terms of number of scientific documents in Biomedical engineering are: America (36.70%, 4427 documents), China (8.21%, 990 documents), Germany (7.21%, 869 documents), Canada (6.79%, 818 documents), and

England (6.26%, 704 documents). Also, China and Japan from Asia have been placed among ten superior countries of the world.

Table 3. Types of world's biomedical engineering documentsin per year according to source

Year / *Types of records*	*Article*	*Article; proceedings paper*	*Review*	*Review; book chapter*	*Meeting abstract*	*Frequency of records according to year*	*Percentage of records according to year*
2002	619	40	20	0	0	679	5.63
2003	726	46	23	0	0	795	6.60
2004	676	111	32	0	0	819	6.80
2005	757	97	27	3	13	897	7.44
2006	855	87	61	4	1	1008	8.36
2007	1074	61	69	6	1	1211	10.05
2008	1253	79	54	6	13	1405	11.66
2009	1481	47	74	5	6	1613	13.39
2010	1729	39	65	7	5	1845	15.31
2011	1652	30	73	6	11	1772	14.71
Total	10822	637	498	37	50	12044	100%
Percent	89.85	5.28	4.13	0.30	0.41		

Based on the statistics presented in Table 3, 89/80% of documents (10822 documents) are articles and 0.3% of them (37 documents) are book reviews. Also, other data of the Table show that the number of scientific productions in Biomedical engineering had an ascending growth up to 2010 and with a fall about 0.6% reached to 1772 documents in the year 2011.

Table 4. Status of top subjects areas according to co-authorship

			Status of co-authorship				
Subjects	*Frequency of records*	*Percentage of Records*	*Single author*	*2author*	*3 author*	*4author*	*Over 4 author*
Materials Science	5896	48.95	75	544	871	1035	3371
Biophysics	1904	15.80	88	334	450	407	625
Sports Science	863	7.16	27	122	195	186	333
Computer science	648	5.38	29	100	141	134	244
Orthopedics	637	5.28	15	70	125	134	293
Medical information	408	3.38	37	53	90	77	151
Polymer Science	297	2.46	5	38	58	46	150
Transplantation	296	2.45	33	23	35	42	163
Radiology, medical imaging and nuclear medicine	229	1.90	4	36	33	42	114
Mathematical Computational Biology	213	1.76	10	33	54	46	70
Total number of authors			323	1353	2052	2149	5514

The status of Medical engineering's highly-worked topics based on multi-authorshipIn Table 4, topics related to medical engineering area and its sub-topics based on the number of their authors have been presented. By investigating the Science Citation Index, 32 topics with frequency of 12011 are identifiable for medical engineering which due to their large number; only ten first topics are referred to.

Table 4 presents topics for medical engineering documents based on multi-authorship. The most percentage of the records belong to material science (48.95%, 5896 documents), biophysics (15.80%, 1904 documents), sports sciences (7.16%, 863 documents), computer (5.38%, 648 documents), orthopedics (5.28%, 737 documents).

Also the Table's data show that all the ten active topics in medical engineering area have more than 4 authors. Therefore, the status of Medical engineering's highly-worked topics based on multi-authorship shows that these topics have more than 4 authors (5514), 4 authors (2149), 3 authors (2052), 2 authors (1353), and 1 author (323), respectively.

Table 5. Cooperator countries with the world's top countries

Country	*Number of cooperator countries*	*Total Common cooperation*	*cooperator countries in order to number of common documents*				
			First colleague	*Second colleague*	*A third colleague*	*Fourth colleague*	*Fifth colleague*
U.S.A.	62	1207	China (121)	Canada (110)	Germany (93)	Japan (78)	England (63)
Germany	46	484	U.S.A. (93)	Switzerland (61)	England & Italy, each (33)	Canada (25)	Austria (21)
England	45	478	U.S.A. (63)	Italy (41)	Germany & Scotland, each (33)	Canada (25)	China (24)
China	30	341	U.S.A. (121)	Japan (32)	Australia & England, each (24)	Singapore (20)	South Korea (17)
Canada	40	335	U.S.A.(110)	France (37)	England & Germany, each (25)	Switzerland (21)	China (14)
Italy	35	310	U.S.A. (57)	England (41)	Germany (33)	France & Spain, each (21)	Switzerland (20)
Switzer-land	35	294	Germany (61)	U.S.A. (57)	England (23)	Canada (21)	Italy (20)
France	38	286	U.S.A. (53)	Canada (37)	England & Italy, each (21)	Switzerland (18)	Spain (17)
Japan	36	236	U.S.A. (78)	China(32)	South Korea (24)	Portugal (10)	England (9)
Spain	33	212	U.S.A. (40)	Portugal(25)	Italy (21)	France (17)	England & Germany, each (14)

In Table 5, the names of world's ten important countries in terms of the number of scientific productions in medical engineering along with their first five peers are presented. Those countries that have more common documents with the intended country are identified as the first peer, and the 2nd- 5th peer countries have less common documents, respectively. All in all, America with 1207 cooperation (62 countries) is in the first place and then, Germany (484, 46 countries), England (478, 45 countries), China (341, 30 countries), Canada (335, 40 countries) are in the second till fourth places regarding the amount of cooperation in scientific productions. America itself has had the highest scientific cooperation with China (121).

Table 6. Frequency of biomedical engineering outputs from world's top ten authors & their collaboration & citations

Author	*Number of scientific outputs*					*Total scientific outputs*	*Percentage of scientific outputs*	*Number of citations to author*	*Author's country*
	Single author	*2 author*	*3 author*	*4 author*	*Over 4 author*				
Kaplan DL	0	0	6	9	67	82	0.68	4802	U.S.A.
Reis RL	0	5	18	12	47	82	0.68	1853	Portugal
Jansen JA	0	1	1	9	34	45	0.37	1238	Netherlands
Sacks MS	0	4	12	13	16	45	0.37	1194	U.S.A.
Athanasiou KA	0	19	13	1	7	40	0.33	785	U.S.A.
Anderson JM	1	3	7	10	17	38	0.31	669	U.S.A.
An KN	0	0	5	7	24	36	0.298	278	U.S.A.
Giardino R	0	0	0	4	31	35	0.290	577	Italy
Langer R	0	1	2	2	30	35	0.290	1906	U.S.A.
Gefen A	9	5	5	11	4	34	0.282	580	Israel
Fini M	0	0	0	4	29	33	0.273	575	Italy
Viceconti M	0	1	1	6	25	33	0.273	419	Italy
Kirkpatrick CJ	0	0	2	3	26	31	0.257	758	Germany
Mano JF	0	1	7	3	20	31	0.257	514	Portugal
Doblare M	0	1	4	9	16	30	0.249	395	Spain
Humphrey JD	1	9	8	6	6	30	0.249	506	U.S.A.
Ratner BD	1	1	2	5	21	30	0.249	1139	U.S.A.
Wang J	0	0	1	5	24	30	0.249	450	China
Zhang Y	0	4	4	4	18	30	0.249	657	China
Chen GQ	0	2	8	4	15	29	0.240	973	China

According to Table 6, the most active authors (8 out of 20, authors, 40%)are from America. Italy, China and Portugal are in the next places. In other words, 8 authors from America,8 authors from Europe and 4 authors from Asia are included in the world's ten active authors in terms of scientific productions in medical engineering major. Looking at Table 6, it can be found that active authors are inclined toward group research work and the number of their 0ne-author productions is zero or one document, except one author from Israel. Other data of this Table shows that Kaplan DL and Reis RL with 82 scientific productions (0.68%) are the first active authors in medical engineering field, and then Jansen JA and Sacks MS with 45 scientific productions (0.37%) are in the second place. Athanasiou KA with 40 scientific productions (0.33%) is the third, Anderson JM with 38 scientific productions (0.31%) is the fourth, and An KN with 24 scientific productions (0.33%) is the fifth person in the world's scientific productions in medical engineering. The amount of references to authors shows that Kaplan DL with 4802 references is in the first place. Langer R who is not among the first five active authors, is the second author with highest references (1906). Reis RL, Jansen JA and Sacks MS with 1853, 1238, and 1194 references are third to fifth highly-referred authors, respectively.

Table 7. The world's top institutions in biomedical engineering scientific outputs

Institution	*Country*	*Number of scientific outputs*	*Percent*
UnivCalif System	Usa	310	2.57
Univ Montreal	Canada	213	1.76
Harvard Univ	Usa	185	1.53
Univ Toronto	Canada	157	1.30
Univ Pittsburgh	Usa	154	1.27
MIT[1]	Usa	148	1.228
Georgia InstTechnol	Usa	147	1.220
UnivTecn Lisbon	Portugal	143	1.18
Univ Michigan	Usa	132	1.09
Univ Bologna	Italy	126	1.04
Total		1715	14.23

The results of Table 7 show that the world's most active institutions and universities after America are from Canada, Portugal, and Italy. Active institutions are first Calif System University with 310 scientific productions, then Montreal University, Harvard University, Toronto University and Pittsburgh University with 213, 185, 157 and 154, respectively.

World's superior countries' co-authorship network

The first 20 countries in Science Citation Index which had more scientific productions in medical engineering field were chosen for illustrating the co-authorship network. However, the information for the first 10 countries is presented in Table 8 which is about the network's information. The network's density is 0.673684211. This fairly high density implies the high links among the nodes. In Table 8 which is calculated by NodeXL software, the status of the world's superior countries is regarded from different aspects.

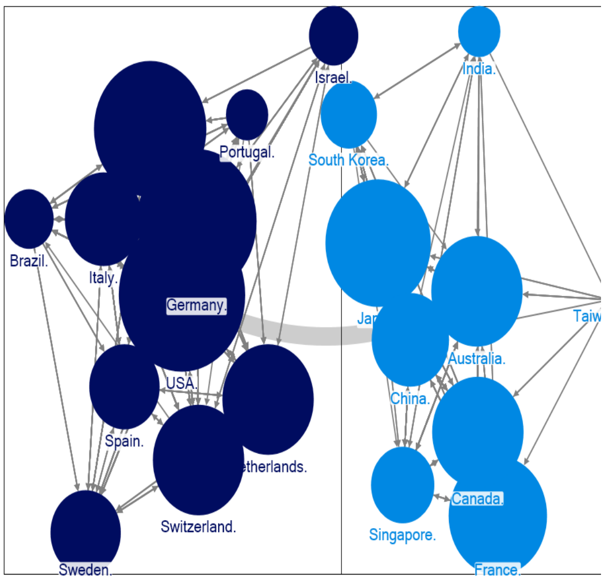


Figure 1. Co-authorship network's clustering of the world's top countries

Table 8. The world's top ten nodes according to various positions in co-authorship networks

[1] Massachusetts Institute of Technology

Ten top nodes according to Out Degree	Out Degree	Ten top nodes according to In Degree	In Degree
USA.	19	USA.	19
Canada.	18	Germany.	18
Germany.	17	England.	18
France.	17	Japan.	17
Japan.	16	France.	16
England.	15	Canada.	15
Italy.	15	Switzerland.	15
Switzerland.	14	Australia	15
China.	14	Netherlands.	15
Netherlands.	13	Italy.	13

Ten top nodes according to Closeness Centrality	Closeness Centrality	Ten top nodes according to Betweenness Centrality	Betweenness Centrality
USA.	0.052632	Canada.	9.991314
Canada.	0.052632	USA.	9.991314
Germany.	0.052632	Germany.	9.991314
England.	0.05	France.	8.654928
France.	0.05	Japan.	7.769886
Japan.	0.05	England.	6.548457
Switzerland.	0.047619	Switzerland.	5.788933
Australia.	0.045455	Australia.	5.086386
Italy.	0.043478	China.	5.030852

Ten top nodes according to Clustering Coefficient	Clustering Coefficient	Ten top nodes according to Eigen Vector Centrality	Eigen Vector Centrality
Brazil.	0.936363636	Canada.	0.061165
Israel.	0.922222222	Germany.	0.061165
Taiwan.	0.892857143	USA.	0.061165
Portugal.	0.83974359	England.	0.059391
Spain.	0.83974359	Japan.	0.058946
Singapore.	0.83974359	France.	0.058396
Sweden.	0.82967033	Switzerland.	0.05637
India.	0.818181818	Australia.	0.053207
Netherlands.	0.804761905	Netherlands.	0.051651

According to Table 8, the in degree information show that America, Germany, England, Japan and France are referred the most by other countries. In other words, these countries have been chosen as the first option for scientific cooperation by other countries. The out degree information shows that compared to other countries, America, Canada, Germany, France and Japan have felt more need for cooperation and have used other countries' help in their scientific productions. Also, the betweenness centrality shows that Canada, America, Germany, France and Japan have the highest status in the network and their non-existence leads to a disorder in the relationships between other countries. As can be seen, although America had the first place in the two previous Tables, in the betweenness centrality Table is placed in the second rank. It implies that degree centrality and betweenness centrality are two distinct categories and high interaction of one node with other nodes has nothing to do with increase in the status and fame of that node in the network. With regard to the closeness centrality, Taiwan, Israel, Brazil, India, South Korea, Portugal, Spain and Singapore from the bottom of table are more influential in distributing the network's information. That is, their average distance from other nodes is less than others and put it simply, they are more available. Interestingly, the countries located at the top of the Table for superior countries, although having a better status in the network, are not so close to other nodes and therefore, have a higher closeness centrality measure. Superior countries according to Eigen Vector centrality also imply that Canada, Germany, America, England, Japan and France more than others are connected with active countries that have a higher degree and more salient social role. With regard to the clustering coefficient, there is a fine point to mention. Those countries that have been in the middle or bottom of the previous tables now have come to the top of this table. It indicates that weaker countries are more inclined toward making more connections with others and establishing scientific groups around their own nodes. In the other hand, the number of links among their neighbors is also more and it increases their clustering coefficient. Brazil, Israel, Taiwan, Portugal, Spain and Singapore have higher clustering coefficient, respectively.

Discussion and conclusion

The research findings show that Europe and Africa has had the most and the least scientific productions in medical engineering in the world, respectively. America, China, Germany, Canada and England were identifies as the world's superior countries in medical engineering outputs. Also, China and Japan from Asia were placed among the world's ten superior countries. The procedural growth of scientific productions in this area had an ascending growth till the year 2010, and then in the year 2011 had a fall for about 6%.

About 90% of documents are articles and the least amount is related to book reviews. Regarding the medical engineering's topics based on multi-authorship, the results show that all the first ten topics have more than 4 authors. Also, most of the topics are about material sciences, Biophysics, sport sciences, computer and orthopedics.

America, Germany, England, China, Canada, Italy, Switzerland, France, Japan and Spain have first to tenth ranks in cooperation in medical engineering's scientific productions. Most of active authors have been from America, Italy, China and Portugal. 8 authors from America continent, 8 authors from Europe, and 4 authors from Asia wereamong the world's 20 active authors.Kaplan DL from America, Reis RL from Portugal, and Jansen Ja from Netherland were the world's 3 active authors. Also, active authors have devoted the highest amount of references to themselves.

The world's most active institutions and Universities in the area of medical engineering's scientific productions after America are from Canada, Portugal and Italy. The world's five superior research centers are: America's Calif system University, Canada's Montreal University, America's Harvard University, Canada's Toronto University and America's Pittsburgh University.

In co-authorship network of the world's 20 superior countries in the field of Biomedical engineering, the United States was the greatest node which had the most connections with other countries. Furthermore, there were two distinct clusters in the network: the first cluster's countries except Israel were from Europe and America continents, and the second cluster's countries were mainly from Asia. The network's rather high density (0.673) implied the existence of a lot of links between nodes. America, Germany, England, Japan and France had the highest amount of in degree. In other words, these countries have been chosen as the first options for scientific cooperation by other countries. On the other hand, America, Canada, Germany, France and Japan more than other countries have felt the need to cooperate with others and have used their help in their scientific productions and had a more out degree.

Canada, America, Germany, France and Japan had the highest status in the network and the higher betweens centrality. Regarding closeness centrality, Taiwan, Israel, Brazil, India, South Korea, Portugal, Spain, and Singapore from the bottom of the table were more effective and more available in the network information publication. Canada, Germany, America, England, Japan and France had a higher Eigen vector. Another important point is the high tendency of weaker countries to connect more with others and make scientific groups around their own node and it has raised their clustering coefficient. Brazil, Israel, Taiwan, Portugal, Spain and Singapore had the higher clustering coefficient, respectively.

All in all, the research findings show that the growth of medical engineering's scientific productions was higher among industrial countries. In ISI Table, superior countries in terms of the amount of scientific productions have had a good productive and industrial power and growth. The best authors, institutions and Universities are from the world's developed and industrial countries. The most productive countries were those which had more scientific cooperation with other countries. Another important point is that high-productive authors are more inclined towards group work; in addition, group work leads to the increase in researchers' scientific productions. The countries of the world's scientific pole which have significant developments in all the areas, have a more significant presence in co-authorship networks and more than others are invited to scientific projects. The importance of the aforementioned points illustrates the necessity and role of the governments' planning and policy-making for establishing the relationships and developing the scientific cooperation among them.

References

Ardanuy, J. (2011). Scientific collaboration in library and information science viewed through the web of knowledge: the spanish case. *Scientometrics*, 90 (3), 877-890.

Bookestein, A. (1994). Scientometrics: New opportunities. *Scientometrics*, 30 (2-3), 455-460.

Center of excellence for biomedical engineering (2007). *Scientific and technology book month*, 1 (4), 38-42.

Cheong, F. &Corbitt, B.J. (2009).A social network analysis of the co-authorship network of the Pacific Asia Conference on Information Systems from 1993 to 2008.*Pacific Asia Conference on Information Systems* (PACIS).

Hasanzadeh, M. &Baghaee, S. (2009). Scientific society, scientific relations & co-authorship.*Rahyaft,* (37), 41-44.

He, T. (2009).International scientific collaboration of China with the G7 countries.*Scientometrics*, 80 (3),571-582.

Khosroabadi, R.&ZamaniNezhad, S. (2007). *Biomedical engineering: principles of operation and maintenance of medical equipment*. Tehran:Modaresanesharif Publications.

Mashhad university of biomedical engineering. Retrieved june 27, 2012 from:http://mbme.ir.

Moody, J. (2004). The structure of a social science collaboration network: Disciplinary cohesion from 1963-1999. *American Sociological Review*, 69(2), 213-238.

Nikzad, M. (2011).Comparative study of co-authorship networks of Iranian articles in library and information science, psychology, management and economics in ISI during 2000-2009 (Master's Thesis). Islamic Azad University, Science and Research Branch.

Osareh,F., NorouziChakoli, A.&Keshvari, M. (2010). Co-authorship of iranian researchers in science, social science, art and humanities citation indexes in the web of science between 2000 and 2006.*Information science & technology*, 25 (4), 573-595.

Teodorescu, D.& Andrei, T. (2011). The growth of international collaboration in East European scholarly communities: A bibliometric analysis of journal articles published between 1989 and 2009. *Scientometrics,* 89, 711-722.

Velayati, K.(2008).Evaluation of scientific cooperation between Iran and neighboring countries from 1990 to 2007 (Master's Thesis).Tehran University, Faculty of Psychology and Educational Sciences.

Ordonez-Matamoros, G., Cozzens, S.E.& Garcia-Luque, M. [2010?].*International co-authorship and research team performance in Colombia*. [Colombia?]: Georgia Institute of Technology.

Women in Science and Higher Education: a bibliometric study

Tahereh Dehdarirad*, Anna Villarroya** and Maite Barrios***

*Department of Library and Information Science, University of Barcelona, Melcior de Palau, 140, 08014 Barcelona (Spain)
tdehdari@gmail.com

**Department of Public Economy, Political Economy and Spanish Economy, University of Barcelona, Melcior de Palau, 140, 08014 Barcelona (Spain)
annavillarroya@ub.edu

***Department of Methodology of Behavioral Sciences, University of Barcelona, Melcior de Palau, 140, 08014 Barcelona (Spain)
mbarrios@ub.edu

Abstract

The main aim of this paper is to study the evolution of the scientific literature on the participation of women in science and higher education. A total of 1,415 articles and reviews published between 1991 and 2012 were extracted from the Thomson Reuters Web of Science database. Standard bibliometric indicators and laws were applied to these data. Furthermore, for each country, Gender Inequality Index (GII) values were calculated in order to produce an international ranking. The results suggest an upward trend in the number of papers. Although the data show low levels of international collaboration, there has been a slight increase in recent years. The interest in gender differences in science and higher education has extended to many different countries (n = 67) and to many scientific journals (n = 595). A Bradford's law analysis revealed a high dispersion of the literature and also a small set of core journals focused on the topic.

Introduction

The entry and progression of women in scientific teaching and research has led to a more balanced representation of the sexes in all fields of science and at all stages of the academic career. However, women still lack equal opportunities in terms of access to top-level positions, decision-making concerning scientific policies and research, and funding. As a consequence, numerous reports and initiatives have emerged in different parts of the world designed to analyse the presence of women in science and higher education, and also to call for a more gender-balanced structure of science (League of European Research Universities [LERU], 2012; Deloitte Consulting, 2013; European Commission, 2013; Organization for Economic Co-operation and Development [OECD], 2013).

There is also a large body of literature on different aspects of the participation and performance of men and women in science and higher education. Many of these studies have examined gender disparities in terms of publication productivity, addressing issues such as the number of publications, citations, impact of researchers' output, and patterns of collaboration. In terms of the number of publications, many studies have demonstrated that female academics publish less, on average, than their male colleagues (Symonds, Gemmell, Braisher, Gorringe & Elgar, 2006; Sidhu et al., 2009; Jagsi et al., 2011). However, other reports suggest that there are no significant differences in productivity between the sexes (Lewison, 2001; Mauleón, Bordons & Oppenheim, 2008). Some of these studies have also evaluated the quality of publications, through either the number of citations or the journal impact factor. The results of the literature in this area vary widely; while some studies report no differences in the citation patterns of male and female academics (Ledin, Bornmann, Gannon & Wallon, 2007; Copenheaver, Goldbeck & Cherubini, 2010), others suggest a predominance of citations for female-authored papers (Long, 1992; Symonds et al., 2006; Borrego, Barrios, Villarroya & Ollé, 2010), and others a predominance of citations for papers authored by men (Hunter & Leahey, 2010; Larivière, Vignola-Gagné, Villeneuve, Gélinas & Gingras, 2011). Studies on the journal impact factor have also produced mixed findings: while some have highlighted the similarity of the journals in which women and men publish (Lewison, 2001; Bordons et al., 2003; Mauleón & Bordons, 2006; Mauleón et al., 2008), others have shown that men choose to publish in journals with a higher impact factor (Hunter & Leahey, 2010), or alternatively that it is women who tend to publish in higher impact journals (Borrego et al., 2010). Another approach to the study of scientific activity concerns the collaborative practices of researchers. Here, there is ample evidence showing that women collaborate to a lesser extent with foreign authors than men (Lemoine, 1992; Lewison, 2001; Webster, 2001; Larivière et al., 2011).

Another sizeable body of research on women in science and higher education has sought to document disparities in academic activities such as grant and manuscript reviewing, obtaining access to funding, and career progression, among others. With regard to manuscript reviewing, Budden et al. (2008) reported that the acceptance rate for female first-authored manuscripts had increased since the advent of blind review. However, further work on this issue has found no differences in the acceptance/rejection ratio for papers submitted by male and female corresponding authors (Aarssen et al., 2008). With regard to gaining access to funding, some studies suggest that male scientists face fewer difficulties in obtaining financial support and better facilities (Stack, 2004; Larivière et al., 2011; LERU, 2012), since they are more likely to hold high status positions from which it is possible to apply for and receive larger grants (Blake & La Valle, 2000; Waisbren et al., 2008). In terms of career progression, many studies have mentioned bias in hiring (Isaac, Lee & Carnes, 2009), the overrepresentation of women in lower faculty ranks (D'Amico, Vermigli & Canetto, 2011) due to the difficulties of ascending the academic ladder (LERU, 2012), and gender pay gaps (Ward, 2001).

These systematic disparities have been attributed to a variety of factors. The literature identifies family formation and childrearing among the major causes of female underrepresentation in academia (Sax, Hagedorn, Arredondo & Dicrisi III, 2002; Prozesky, 2008; Hunter & Leahy, 2010). Other factors which have been addressed in different studies are personal and institutional (structural) factors (Dewandre, 2002; Shen, 2013), professional issues (Allison & Long, 1990; Sonnert, 1996; Zinovyeva & Bagues, 2011), psychological and individual issues (Sonnert, 1996) and the level of specialization (Leahey, 2006).

Despite the relevance of the subject, and the number of academic publications, initiatives, and reports on *women in science and higher education*, no systematic analysis has yet been carried out of the large body of research in this area. Using standard bibliometric indicators (such as the number of publications and productivity by country, among others) and laws (Price's and Bradford's laws) this article aims to assess the development and growth of research in this field by reviewing the related scientific literature.

Methodology

Data collection

The data were extracted from the Thomson Reuters Web of Science in February 2013, searching the topic fields from 1991 to 2012. Reviewing the related scientific literature, we noted that studies on women in science and higher education frequently address three main topics: *publication productivity*, *issues related to gender in academia and science* and *factors related to gender bias*. The study of *publication productivity* includes papers dealing with scientific productivity, citation and collaboration patterns. *Issues related to gender in academia and science* refers to papers addressing elements that have an influence on the development of the scientific career such as interviewing and hiring, salaries, promotion and advancement, access to funding, mentoring and networking, and being a member of editorial board or a peer reviewer. Finally, *factors related to gender bias* covers papers assessing the issues that have an incidence on researchers' performance under the two previous headings such as family-related issues and structural, institutional, professional, biological, psychological, social, and political variables. We then conducted three different searches, one for each topic. After elimination of duplicates, a total of 1,225 records were considered. Additionally, in order to ensure that all references dealing with the subject were included in the database, the most recent papers were checked and 190 new papers were added. As a result, a corpus of 1,415 articles and reviews were finally considered. All these papers were coded according to the three headings mentioned above. It should be noted that a paper may simultaneously address more than one issue, and will therefore belong to more than one group.

Data analysis

The main bibliometric laws were applied to study scientific growth over time and the dispersion of scientific output across journals.

Scientific growth over time was assessed using Price's law (Price, 1963). In order to test whether our data followed Price's law, different regression models were fitted, including linear, exponential and logistic curves. The latter were applied to assess the hypothesis of literature growth saturation.

Bradford's law (Bradford 1934, 1948) was used to study the dispersion of the literature. Specifically, Bradford's law describes how the articles in a specific area are distributed across journals, postulating a model of concentric productivity zones with a decreasing information density. Following the proposal of Egghe (1986, 1990), the Bradford multiplier was obtained by $K = (1.781 \cdot y_m)^{1/P}$ in which is y_m the number of articles published by the most productive journal and *P* is the number of zones including the core. The estimated *k* value for each zone was calculated by the ratio between the number of journals in a given zone and the number of journals in any immediate zone. The number of Bradford zones was determined by the solution that minimized the difference between the Bradford multiplier *k* and each estimated value of *k*, and between the estimated values of *k*. In addition, the predicted frequencies were fitted according to Leimkuhler's formulation (Leimkuhler, 1967), obtaining the constants as $A = y_o / \log_e k$ and $B = (k-1)/r_0$ in which y_0 *is* the constant number of articles (in each group where *a is* the total number of articles and both *P* and *k* are as defined above) and r_0 is the expected number of journals in the core ($r_0 = \frac{T(K-1)}{k^p - 1}$ in which *T* is the total number of journals, and *k* and *P* are as defined above). The estimated cumulative number of articles produced by the journals of rank 1, 2…r was obtained by $R(r) = A \cdot \log_e(1 + B \cdot r)$.

Additionally, and for each country, the most recent Gender Inequality Index (GII) was considered in order to rank countries in terms of inequalities. GII is a new index for measuring gender disparity which was introduced in the 2010 Human Development Report (20th anniversary edition) of the United Nations Development Program (UNDP). The UNDP describes this index as a composite measure which captures the loss of achievement within a country due to gender inequality, and it uses three dimensions to do so: reproductive health, empowerment, and labor market participation (United Nations Development Program, 2013).

Results

Number of publications

Figure 1 shows an upward trend in the percentage of publications. Linear, exponential and logistic regression models were fitted in order to test whether the data conformed to Price's law. The exponential fit (R^2= 0.834) showed a higher proportion of the explained variance compared to the linear (R^2= 0.707) or logistic (R^2= 0.578) approaches, presenting a good adjustment to Price's law. The three main topics also showed a good fit to the exponential model. The topic with the highest frequency of papers was *issues related to gender in academia and science* (943, 66.64%), followed by *factors related to gender bias* (438, 30.95%) and *publication productivity* (275, 19.43%).

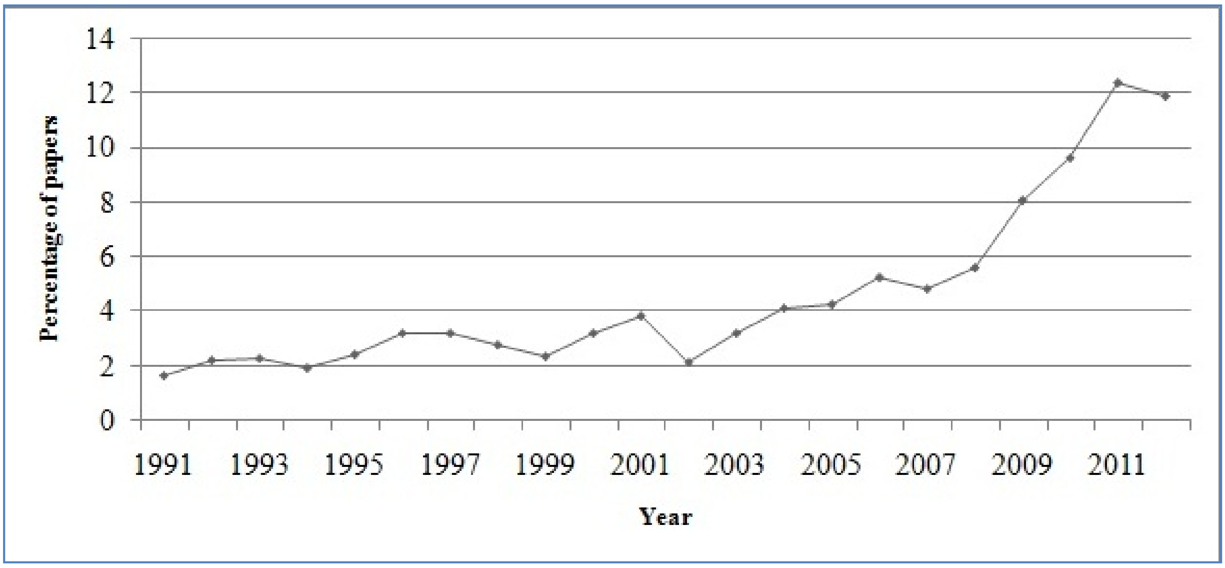


Figure 1. Temporal evolution of percentage of publications

Countries

Sixty-seven countries participated in the data set, although it should be noted that authors' affiliations were not available for 204 papers (14.4%). The top ten countries in terms of contributions were the US, the UK, Canada, Australia, Spain, Germany, the Netherlands, Sweden, China, Brazil, Italy and Turkey (the last three with the same number of publications). Only 8% (n = 113) of papers involved international collaboration, and of these, 22.12% (n = 25) were published in 2012. Table 1 shows the top ten countries in terms of contributions, as well as the most recent GII and regression fit for each country. No statistically significant relationship was found between the recent GII and the number of papers published by each country (r = -.099, p = .449). However, a significant correlation coefficient was obtained between the most recent GII and the number of papers published through international collaboration by each country, controlling for the total number of papers published by each country ($r_{xy.z}$ = -.294, p = .022).

Table 1. Regression fit of countries, publication frequency, and corresponding recent GII

Country	Frequency (%)	Recent GII(2012)	R^2linear	R^2exponential	R^2logistic
USA	638 (45.08)	0.256	0.562	**0.643**	0.521
UK	126 (8.90)	0.205	0.555	0.512	0.513
Canada	70 (4.94)	0.119	0.549	0.**614**	0.575
Australia	63 (4.45)	0.115	0.304	0.349	0.341
Spain	57 (4.02)	0.103	0.568	**0.683**	0.579
Germany	41 (2.89)	0.075	0.336	0.291	0.313
Netherlands	29 (2.04)	0.045	0.519	**0.630**	0.566
Sweden	29 (2.04)	0.055	0.492	0.507	0.496
China	20 (1.41)	0.213	0.288	0.340	0.297
Brazil	17 (1.20)	0.447	0.301	0.325	0.306
Italy	17 (1.20)	0.094	0.219	0.178	0.214
Turkey	17 (1.20)	0.366	0.188	0.158	0.188

Among the papers involving international collaboration, 56.64% (n = 64) dealt with the topic of *issues related to gender in academia and science*, 30.08% (n = 34) addressed *factors related to gender bias*, and 28.31% (n = 32) examined *publication productivity*. As any given paper may simultaneously address more than one topic, the sum of papers is more than the total number of papers, and the sum of percentages exceeds 100%.

Journals

The papers were published in a total of 595 journals. Three hundred and sixty-six (61.5%) journals published only one paper. The distribution of papers published in the set of journals was described using Bradford's law, which revealed that the papers were distributed in four zones and that the core comprised 13 journals. Table 2 shows the expected number of journals given the Bradford multiplier (3.17), the actual number of journals in each zone, the number of articles included in each zone, the cumulative number of articles, the estimated values of *k*, and, finally, the predicted cumulative number of articles R(r). Table 3 shows the core journals and their publication frequency.

Table 2. Data fit to Bradford's law

Zone	*Expected number of journals*	*Number of journals*	*Number of articles*	*Cumulative articles*	*Estimated k*	*R(r)*
1	13	13	366	366	_	355.86
2	41.10	41	250	616	3.15	634.02
3	129.98	130	343	959	3.17	959.77
4	411	411	456	1415	3.16	1303.04

Constants according to Leimkuhler's formulation were A = 306.26 and B = 0.168.

Table 3. Core journals and their publication frequency

Core journal	Frequency (%)
Scientometrics	57 (15.57)
Sex roles	42 (11.48)
Academic medicine	39 (10.66)
Higher Education	36 (9.84)
Research in higher education	31 (8.47)
Gender and education	30 (8.20)
Scientist	29 (7.92)
Women's studies international forum	19 (5.19)
Gender work and organization	18 (4.92)
Journal of higher education	17 (4.64)
Journal of vocational behavior	17 (4.64)
Journal of womens' health	16 (4.37)
Academic psychiatry	15 (4.10)

Discussion and Conclusion

The present study has analysed the main bibliometric indicators in relation to the literature on women's participation in science and higher education. With regard to the number of publications, results showed a significant increase and interest in the field over the last 21 years, particularly since 2002. This increase was supported by the fit of the data to Price's law, which indicates that productivity in the studied field presents exponential growth. Of the three topics considered, namely *publication productivity*, *issues related to gender in academia and science*, and *factors related to gender bias*, the highest number of papers corresponded to the second topic (i.e., *issues related to gender in academia and science*), accounting for 66.64% of the total publications. This topic was also the most frequently addressed in papers involving international collaboration (56.64%), possibly because it addresses a wide variety of issues such as pursuing a scientific career, having access to funding, mentoring and networking, and being a member of an editorial board or a peer reviewer. Furthermore, these studies often consider the other two topics in their attempt to explain the differences in academia and science and use bibliometric indicators such as publication productivity. As for the GII, the results showed a modest relationship between its most recent version and the number of papers published in international collaboration by each country. This means that countries with a lower GII are more likely to collaborate internationally.

Although a slight increase in international collaboration has been reported in recent years, the rate remains very low. These results indicate that this field of study has yet to become truly international and that collaboration between countries and institutions needs to be reinforced.

Sixty-seven countries contribute to the growth of research, with the US and the UK heading the list. The fact that a large number of countries contribute to this field suggests that gender inequality remains a global problem, despite the substantial initiatives and policies undertaken at national and international level. The finding that the US and the UK were the most productive countries may be related to the database used for the analysis. Most of the journals included in this database are written in English and, consequently, they are more likely to originate from either the US or the UK. In fact, several studies have indicated that journals based in the US and the UK are significantly over-represented in the Web of Science, a database that only includes a very limited number of journals in languages other than English (Archambault & Gagné, 2004; Yang & Meho, 2006; Harzing, 2010). This limited coverage, which constitutes a bias in favor of English language journals from English-speaking countries, is therefore a limitation of the present study.

Finally, regarding Bradford's law and the distribution of journals, we identified a small set of core journals in which a large number of papers were concentrated. Scientometrics is the journal with the higher number of papers. This small set of core journals shows the multidisciplinary nature of the topic, since a range of research areas are represented (e.g. Computer Science, Information Science & Library Science, Psychology; Women's Studies, Education & Educational Research; Health Care Sciences & Services).

To sum up, the scientific literature on the participation of women in science and higher education has expanded considerably in recent decades. Scientific journals in many countries are now interested in the field, which encompasses a range of research areas. In the European context, the promotion of gender equality, including the integration of the gender dimension in research and innovation enshrined in the "Horizon 2020", is a clear sign of the current relevance of this topic in the scientific arena and will markedly increase productivity in this area in the near future.

References

Aarssen, L., Tregenza, T., Budden, A. E., Lortie, C. J., Koricheva, J., & Leimu, R. (2008). Bang for your buck: rejection rates and impact factors in ecological journals. *The Open Ecology Journal,* 1, 114-19.doi: 10.2174/1874213000801010014.

Allison, P. D., & Long, J. S. (1990). Departmental effects on scientific productivity. *American Sociological Review*, 55(4), 469-478. doi: 10.2307/2095801.

Archambault, É., & Gagné, É. V. (2004). The use of bibliometrics in the Social Sciences and Humanities *Science-Metrix*. Montreal, Canada: Social Sciences and Humanities Research Council of Canada (SSHRCC).

Blake, M., & La Valle, I. (2000). *Key factors shaping funding application behaviour among women and men in British higher education institutions*. London, UK: Wellcome Trust.

Bordons, M., Morillo, F., Fernández, M. T., & Gómez, I. (2003). One step further in the production of bibliometric indicators at the micro level: Differences by gender and professional category of scientists. *Scientometrics,* 57(2), 159-173. doi: 10.1023/A:1024181400646.

Borrego, Á., Barrios, M., Villarroya, A., & Ollé, C. (2010). Scientific output and impact of postdoctoral scientists: a gender perspective. *Scientometrics*, 83(1), 93-101. doi: 10.1007/s11192-009-0025-y.

Bradford, S. C. (1934). Sources of information on specific subjects. *Engineering*, 23(3), 85–88.

Bradford, S. C. (1948). *Documentation*. London: Crosby Lockwood Sons, Ltd.

Budden, A. E., Tregenza, T., Aarssen, L. W., Koricheva, J., Leimu, R., & Lortie, C. J. (2008). Double-blind review favors increased representation of female authors. *Trends Ecol Evol*, 23(1), 4-6.doi: 10.1016/j.tree.2007.07.008.

Copenheaver, C. A., Goldbeck, K., & Cherubini, P. (2010). Lack of gender bias in citation rates of publications by dendrochronologists: What is unique about this discipline? Tree-Ring Research, 66(2), 127-133. doi: 10.3959/2009-10.1.

D'Amico, R., Vermigli, P., & Canetto, S. S. (2011). Publication productivity and career advancement by female and male psychology faculty: The case of Italy. Journal of Diversity in Higher Education, 4(3), 175-184. doi: 10.1037/a0022570.

Deloitte Consulting. (2013). *Researchers' Report 2013*. Brussels: European Commission's Directorate-General for Research and Innovation.

Dewandre, N. (2002). European Strategies for Promoting Women in Science. *Science,* 295(5553), 278-279. doi: 10.1126/science.1063487.

Egghe, L. (1986). The dual of Bradford's law. *Journal of the American Society for Information Science,* 37(4), 246–255.

Egghe, L. (1990). Applications of the theory of Bradford's law to the calculation of Leimkuhler's law and to the completion of bibliographies. *Journal of the American Society for Information Science*, 41(7), 469–492.

European Commission. (2013). *She Figures. Gender in research and innovation*. Brussels: Directorate General for Research and Innovation.

Harzing, A. W. (2010). *The Publish or Perish Book: Your guide to effective and responsible citation analysis*. Melbourne: Tarma Software Research.

Hunter, L. A., & Leahey, E. (2010). Parenting and research productivity: New evidence and methods. *Social Studies of Science*, 40(3), 433-451. doi: 10.1177/0306312709358472.

Isaac, C., Lee, B., & Carnes, M. (2009). Interventions that affect gender bias in hiring: a systematic review. *Acad Med*, 84(10), 1440-1446. doi: 10.1097/ACM.0b013e3181b6ba00.

Jagsi, R., DeCastro, R., Griffith, K. A., Rangarajan, S., Churchill, C., Stewart, A., & Ubel, P. A. (2011). Similarities and differences in the career trajectories of male and female career development award recipients. *Acad Med*, 86(11), 1415-1421. doi: 10.1097/ACM.0b013e3182305aa6.

Larivière, V., Vignola-Gagné, E., Villeneuve, C., Gélinas, P., & Gingras, Y. (2011). Sex differences in research funding, productivity and impact: an analysis of Québec university professors. *Scientometrics*, 87(3), 483-498. doi: 10.1007/s11192-011-0369-y.

Leahey, E. (2006). Gender Differences in productivity: research specialization as a missing link. *Gender & Society*, 20(6), 754-780.doi: 10.1177/0891243206293030.

Ledin, A., Bornmann, L., Gannon, F., & Wallon, G. (2007). A persistent problem. EMBO Rep, 8(11), 982-987. doi: http://www.nature.com/embor/journal/v8/n11/suppinfo/7401109_S1.html.

Leimkuhler, F. F. (1967). The Bradford distribution. *Journal of Documentation*, 23(3), 197–207.

Lemoine, W. (1992). Productivity patterns of men and women scientists in Venezuela. *Scientometrics,* 24(2), 281-295. doi: 10.1007/BF02017912.

LERU (League of European Research Universities). (2012). *Women, research and universities: excellence without gender bias*. Leuven, Belgium: League of European Research Universities.

Lewison, G. (2001). The quantity and quality of female researchers: A bibliometric study of Iceland. *Scientometrics*, 52(1), 29-43. doi: 10.1023/A:1012794810883.

Long, J. S. (1992). Measures of sex differences in scientific productivity. *Social Forces,* 71(1), 159-178. doi: 10.1093/sf/71.1.159.

Mauleón, E., & Bordons, M. (2006). Productivity, impact and publication habits by gender in the area of Materials Science. *Scientometrics,* 66(1), 199-218. doi: 10.1007/s11192-006-0014-3.

Mauleón, E., Bordons, M., & Oppenheim, C. (2008). The effect of gender on research staff success in life sciences in the Spanish National Research Council. *Research Evaluation*, 17(3), 213-225. doi: 10.3152/095820208x331676.

OECD (Organization for Economic Co-operation and Development). (2013). *Education at a Glance 2013: OECD indicators*. Paris: OECD Publishing.

Price, D. J. d. S. (1963). *Little science, big science*. New York: Columbia University Press.

Prozesky, H. (2008). A career-history analysis of gender differences in publication productivity among South African academics. *Science Studies*, 21(2), 47-67.

Sax, L., Hagedorn, L., Arredondo, M., & Dicrisi, F., III. (2002). Faculty research productivity: Exploring the role of gender and family-related factors. *Research in Higher Education*, 43(4), 423-446. doi: 10.1023/A:1015575616285.

Shen, H. (2013). Inequality quantified: Mind the gender gap. *Nature*, 495(7439), 22-24. doi: 10.1038/495022a.

Sidhu, R., Rajashekhar, P., Lavin, V. L., Parry, J., Attwood, J., Holdcroft, A., & Sanders, D. S. (2009). The gender imbalance in academic medicine: a study of female authorship in the United Kingdom. *J R Soc Med*, 102(8), 337-342. doi: 10.1258/jrsm.2009.080378.

Sonnert, G. (1996). Gender equity in science: Still an elusive goal. *Issues in Science and Technology*, 12(2), 53–58.

Stack, S. (2004). Gender, children and research productivity. *Research in Higher Education*, 45(8), 891-920. doi: 10.1007/s11162-004-5953-z.

Symonds, M. R., Gemmell, N. J., Braisher, T. L., Gorringe, K. L., & Elgar, M. A. (2006). Gender differences in publication output: towards an unbiased metric of research performance. *PLoS One*, 1, e127. doi: 10.1371/journal.pone.0000127.

United Nations Development Programme. (2013). Gender Inequality Index (GII). Retrieved 20/2/2014, from http://hdr.undp.org/en/statistics/gii

Waisbren, S. E., Bowles, H., Hasan, T., Zou, K. H., Emans, S. J., Goldberg, C., & Christou, H. (2008). Gender differences in research grant applications and funding outcomes for medical school faculty. *J Womens Health (Larchmt)*, 17(2), 207-214. doi: 10.1089/jwh.2007.0412.

Webster, B. M. (2001). Polish women in science: a bibliometric analysis of Polish science and its publications, 1980–1999. *Research Evaluation,* 10(3), 185-194. doi: 10.3152/147154401781776999.

Ward, M. (2001). The gender salary gap in British academia. *Applied Economics,* 33(13), 1669-1681. doi: 10.1080/00036840010014445.

Yang, K., & Meho, L. I. (2006). Citation Analysis: A Comparison of Google Scholar, Scopus, and Web of Science. *Proceedings of the American Society for Information Science and Technology, 43*(1), 1-15. doi: 10.1002/meet.14504301185.

Zinovyeva, N. & M. Bagues. (2011). *Does gender matter for academic promotion? Evidence from a randomized natural experiment*. IZA Discussion Paper No. 5537, Bonn, Germany.

Pattern of Research & Citations: A Study of Three Central Universities Located in Delhi-India

Tariq Ashraf

Delhi University Library System University of Delhi-Delhi-110007 -India
tariq@south.du.ac.in

1. Introduction

Research productivity in higher education is gaining importance for the past one decade in India. Faculty members of the universities in India have two functions to perform, teaching and research. Research in universities has gained momentum during the past one and half decade, mainly due to support received through funding projects from major government scientific agencies. The research output of the university scientists in the form of research papers in peer-reviewed scholarly journals is being considered as one of the main criteria for assessing the performance of the university scientists and faculty.

Citation analysis is used to evaluate the journals in a field, and the research conducted in a discipline, by a scholar or even of an entire country. Citation analysis is often used to appraise the performance of a researcher and is one important criterion for promotion and tenure. Despite criticism, it has value as an analytical tool, since, as Baird and Oppenheim stated, "whatever measure you take for the eminence of an individual scientist or of a journal or of an institution, citation counts provide strong correlation with that result". They further declared that "high citation counts mean a statistical likelihood of high quality research".

2. Universities In Delhi

The University of Delhi , Jawahar Lal Nehru University and Jamia Millia Islamia are three premier and nationally renowned are universities of the country and are known for their high standards in teaching and research and attracts eminent scholars to their faculty. Ever since their inception, a strong commitment to excellence in teaching and research has made these Universities role-models and path-setter for other universities in the country. Their rich academic tradition has always attracted the most talented students who later on went on to make important contributions to their society.

2.1 Jawahar Lal Nehru University

Established in 1969 by Act of Parliament ,JNU is very prestigious central university designed for interdisciplinary approach in teaching and research with freedom to define and design course content or start new courses. Research themes evolve with new developments in the area and the interface between different areas of study. Everyone at the university competes with himself/herself to excel in their own field of research. JNU is academically and socially a vibrant place where all have space to express their views.

The ten Schools and four special centres of the university produce high quality research publications, books, working papers and M.Phil and Ph,D theses. The JNU alumni occupy important positions, in academics, government, private sector, and in fact in all walks of life. Recently JNU has been ranked by the NAAC as the top University in the country.

2.2 Jamia Millia Islamia

Jamia Millia Islamia came into existence in1920. By a Special Act of the Parliament, Jamia Millia Islamia was made a central university of India in December 1988. In the list of the

Faculties, i.e. Education, Humanities & Languages, Natural Sciences, Social Sciences. Engineering & Technology, one more Faculty - Faculty of Law, was added in 1989. Many new courses and programmes at UG and PG levels have since been added.

Besides its Nine faculties, the Jamia has a number of centres of learning and research, like AJK-Mass Communication Research Centre (MCRC), Academy of International Studies etc. The Jamia is also marching ahead in the field of Information Technology (IT). It offers various undergraduate and postgraduate IT courses.

The paper attempts to understand the pattern of research publications during last one decade of these three universities and understand their relative contribution to research. It also examines the citation pattern of published paper so as to ascertain their quality using Scopus, the citation database.

3. Scopus

Scopus, the largest abstract and citation database of peer-reviewed literature, features smart tools to track, analyze and visualize research. Scopus delivers an overview of the world's research output in the fields of science, technology, medicine, social sciences and arts and humanities. Updated daily, Scopus includes: 21,000 titles from more than 5,000 international publishers 20,000 peer-reviewed journals (including 2,600 open access journals) 390 trade publications 370 book series, 5.5 million conference papers"Articles-in-Press" from more than 3,850 journals and publishers such as Cambridge University Press, Elsevier, Springer, Wiley-Blackwell, Nature Publishing Group and the Institute of Electrical and Electronics Engineers. Scopus track citations over time for a set of authors or documents, with Citation Overview/Tracker and set citation alerts.

4. Review of Related Literature

A literature review is a body of text that aims to review the critical points of current knowledge including substantive findings as well as theoretical and methodological contributions to a particular topic. Literature reviews are secondary sources, and as such, do not report any new or original experimental work. A literature review usually precedes a research proposal and results section. Its ultimate goal is to bring the reader up to date with current literature on a topic and forms the basis for another goal, such as future research that may be needed in the area.

A short review of relevant literature is presented below:

Peters et. al. (1988) used the literature of Chemical Engineering to monitor scientific productivity or "research performance". They found bibliometric analysis to be a valuable tool. Their methodology seemed designed to reinforce rather than examine preconceived notions, however. They suggested that a citation analyses was not useful to the same extent in all sub-fields of chemical engineering since some scientists with "high reputations" turned out to be "bibliometrically invisible". Twenty-one "significant" journals of oceanography were examined by Garfield. (1988) Here too, a concern was raised over sub-fields within the discipline. For the purpose of this study, oceanography meant chemical and physical oceanography. Marine biology was excluded. Garfield's intent here was to prioritize journal titles worldwide using the citation patterns of titles from this "core" list.

The work by Spies (1991) reviewed fourteen "major" journals of exploration geophysics for "effectiveness". This paper presented a basic tenet of citation analysis as follows: the references that an author cites are a roughly valid indicator of influence, hence value, to his work. A measure of cost effectiveness was also incorporated into this study as subscription costs and citation rates were compared between commercial publications and those produced by professional societies.

Uzun et. al. (1993) studied the citation rates of 572 Turkish physics publications that appeared in the source journals listed in the Science Citation Index. This analysis was global in scope, as is commonly the case with citation studies. They examined impact factor, immediacy index, citation frequency, and the nationality of the publishing house. The question which these normal parameters and this normal scope raise for librarians is - "How is this relevant to the local collection for which I have a responsibility to build based upon the research and curriculum program ongoing at my institution"?

Rousseau (1988) presented a citation distribution of mathematics journals, wherein he proposes that a four-year impact factor would be more suited to mathematics than the more or less typical two-year impact factor used in Science Citation Index. Once again, the question of relevancy presents itself.

5. Objectives of the Study

The main objectives of the study are:

- To study the research output of these three universities located in Delhi
- To study the publication pattern of these Universities
- To analyze the contribution of the authors of these three universities.
- To study the authorship pattern and author productivity.
- To identify strong and weak disciplines
- To find out the funding Sources.
- To study the citation pattern of research papers h-Index of journals.
- To find out the relative contribution of each university to research

6. Methodology

The data for study shall be downloaded from the Web of Science database. Web of Science is an international multidisciplinary database indexing over 50 million records. Publications data for twelve years from 2000 to 2014, shall be used for analyzing the research pattern and citation analysis of these universities.

7. Limitation

The study has been conducted using the bibliographic and citation database , Scopus and all the results are based upon the journals and other published material covered by Scopus. The paper does not take into account any research produced by three universities beyond the coverage of Scopus database.

8. Data Analysis and Interpretation

Table 1: Total Documents Output (DU)

Year	Search Field	No. of Records
1935-2014	University of Delhi	17,128

Yearly average of papers published comes to 217 for last 79 years for which data is available.

Table 1.1: Total Documents Output (JMI)

Year	Search Field	No. of Records
1973-2014	Jamia Millia Islamia	1008

Yearly average of papers published comes to 25 for last 41 years for which data is available.

Table 1.2: Total Documents Output (JNU)

Year	Search Field	No. of Records
1969-2014	Jawaharlal Nehru University	6,344

Yearly average of papers published comes to 140 for last 45 years for which data is available.

Table 2: Subject-wise Distribution (DU)

Subject Area	Record Count	% of 17,128
Physics and Astronomy	4961	28.964%
Chemistry	3494	20.399%
Biochemistry, Genetics and Molecular Biology	2752	16.067%
Agricultural and Biological Sciences	2019	11.787%
Materials Science	1951	11.390%
Engineering	1891	11.040%
Mathematics	1455	8.494%
Medicine	1381	8.062%
Pharmacology, Toxicology and Pharmaceutics	1008	5.885%
Computer Science	858	5.009%
Chemical Engineering	781	4.559%
Social Sciences	757	4.490%
Earth and Planetary Sciences	756	4.484%
Environmental Science	717	4.186%
Immunology and Microbiology	566	3.304%
Decision Sciences	428	2.498%
Multidisciplinary	310	1.809%
Economics, Econometrics and Finance	269	1.570%
Business, Management and Accounting	206	1.202%
Undefined	205	1.196%
Energy	191	1.115%
Arts and Humanities	171	0.998%
Psychology	162	0.945%
Neuroscience	70	0.408%
Health Professions	35	0.204%
Nursing	25	0.145%
Veterinary	18	0.105%
Dentistry	1	0.005%

Highest number of papers have been published in the area of Physics and Astronomy followed by Chemistry and Biotechnology, Genetics and Molecular Biology. The least number of papers published are in the area of Dentistry.

Table 2.1: Subject-wise Distribution (JMI)

Subject Area	Record Count	% of 1,766
Chemistry	242	13.703%
Engineering	212	12.004%
Physics and Astronomy	206	11.664%
Materials Science	196	11.098%
Biochemistry, Genetics and Molecular Biology	142	8.040%
Computer Science	138	7.814%
Chemical Engineering	96	5.436%
Mathematics	84	4.756%
Medicine	75	4.246%
Pharmacology, Toxicology and Pharmaceutics	65	3.680%
Environmental Science	55	3.114%
Social Sciences	53	3.001%
Agricultural and Biological Sciences	49	2.774%
Energy	34	1.925%
Immunology and Microbiology	26	1.472%
Business, Management and Accounting	20	1.132%
Earth and Planetary Sciences	18	1.019%
Arts and Humanities	14	0.792%
Economics, Econometrics and Finance	13	0.736%
Multidisciplinary	8	0.453%
Decision Sciences	7	0.396%
Dentistry	4	0.226%
Psychology	3	0.169%
Nursing	2	0.113%
Health Professions	2	0.113%
Undefined	2	0.113%

The subject wise distribution of JMI reveals that highest number of papers have been published are in the area of Chemistry followed by Engineering, Physics and Astronomy.

Table 2.2: Subject-wise Distribution (JNU)

Subject Area	Record Count	% of 6,344
Biochemistry, Genetics and Molecular Biology	1679	26.465%
Social Sciences	1026	16.172%
Physics and Astronomy	940	14.817%
Medicine	847	13.351%
Agricultural and Biological Sciences	811	12.783%
Environmental Science	701	11.049%
Computer Science	441	6.951%
Mathematics	432	6.809%
Immunology and Microbiology	421	6.636%

Economics, Econometrics and Finance	407	6.415%
Earth and Planetary Sciences	405	6.383%
Chemistry	327	5.154%
Engineering	322	5.075%
Materials Science	309	4.870%
Pharmacology, Toxicology and Pharmaceutics	275	4.334%
Chemical Engineering	263	4.145%
Arts and Humanities	175	2.758%
Multidisciplinary	129	2.033%
Neuroscience	126	1.986%
Business, Management and Accounting	98	1.544%
Energy	73	1.150%
Decision Sciences	70	1.103%
Health Professions	56	0.882%
Undefined	33	0.520%
Psychology	30	0.472%
Nursing	28	0.441%
Veterinary	15	0.236%
Dentistry	2	0.0315%

The data reveals that highest number of papers have been published in the interdisciplinary area of Chemistry , Engineering , Physics and Astronomy followed by Social Sciences and Physics and Astronomy. It clearly discernible from data JNU is quite strong in the area of social science research in comparison to University of Delhi.

Table 3: Publication Types (DU)

Field: Documents Type	**Record Count**	**% of 17,128**
Article	14427	84.230%
Conference Paper	1173	6.848%
Review	543	3.170%
Undefined	389	2.271%
Letter	221	1.290%
Article in Press	99	0.578%
Note	82	0.478%
Book Chapter	66	0.385%
Editorial	52	0.303%
Erratum	48	0.280%
Short Survey	23	0.134%
Book	5	0.029%

84.23 percent the publications are in the form of articles which is most established format for research output and followed by conference papers.

Table 3.1: Publication Types (JMI)

Field: Documents Type	Record Count	% of 998
Article	743	73.546%
Conference Paper	166	16.633%
Review	58	5.811%
Article in Press	18	1.803%
Book Chapter	9	0.901%
Letter	5	0.501%
Note	3	0.300%
Erratum	3	0.300%
Short Survey	1	0.100%
Undefined	1	0.100%
Editorial	1	0.100%

The 73.5 percent publications are in the form of articles whereas 16.6 percent research is in the form of conference papers.

Table 3.2: Publication Types (JNU)

Field: Documents Type	Record Count	% of 6,344
Article	5000	78.814%
Conference Paper	444	6.998%
Review	409	6.447%
Undefined	187	2.947%
Book Chapter	64	1.008%
Note	53	0.835%
Article in Press	50	0.788%
Letter	49	0.772%
Editorial	35	0.551%
Erratum	24	0.378%
Short Survey	24	0.378%
Book	5	0.078%

78.8 percent of JNU research has been published in the form of articles whereas 6.99 percent in the form of conference papers.

Table 4: Authorship Pattern (DU)

Field: Authors	Record Count	% of 17,128
Shivpuri, R.K.	687	4.010%
Choi, S.	654	3.818%
Varelas, N.	641	3.742%
Bhatnagar, V.	640	3.736%
Beri, S.B.	638	3.724%
Hirosky, R.	628	3.666%
Cutts, D.	626	3.654%
Wayne, M.	624	3.643%
Gerber, C.E.	623	3.637%
Ruchti, R.	620	3.619%
Banerjee, S.	618	3.608%

Snow, G.R.	612	3.573%
Elvira, V.D.	609	3.555%
Bhat, P.C.	608	3.549%
Barberis, E.	605	3.532%
Zielinski, M.	604	3.526%
Narain, M.	600	3.503%
Gavrilov, V.	598	3.491%
Heintz, U.	595	3.473%
Demina, R.	594	3.468%
Landsberg, G.	590	3.444%
Ellison, J.	585	3.415%
Lipton, R.	575	3.357%
Hagopian, S.	574	3.351%
Gershtein, Y.	573	3.345%

If we look at the authorship pattern we find that topmost author has contributed slightly over percent to the research profile of University of Delhi whereas 25 authors have contributed more than 3 percent. However this percentage is not absolute since most of the authors have contributed jointly.

Table 4.1: Authorship Pattern (JMI)

Field: Authors	**Record Count**	**% of 998**
Ali, I.	72	7.214%
Husain, M.	66	6.613%
Ali, A.	64	6.412%
Khan, Z.	49	4.909%
Zulfequar, M.	29	2.905%
Islam, S.S.	27	2.705%
Aboul-Enein, H.Y.	27	2.705%
Malik, M.A.	26	2.605%
Saleem, K.	25	2.505%
Abulaish, M.	25	2.505%
Fatma, T.	23	2.304%
Dewan, K.K.	20	2.004%
Islam, T.	20	2.004%
Nain, A.K.	20	2.004%
Khan, Z.H.	18	1.803%
Hashmi, A.A.	18	1.803%
Awana, V.P.S.	17	1.703%
Salah, N.	17	1.703%
Habib, S.	17	1.703%
Kishan, H.	15	1.503%
Ikram, S.	15	1.103%
Harsh,	15	1.503%
Tariq, M.	14	1.402%
Zaheeruddin,	14	1.402%
Patel, R.	13	1.302%

The authorship pattern of Jamia Millia Islamia looks less concentrated though top most author has contributed more than 7 percent of total research followed by two authors contributing 6 percent or so each. The percentage get down to nearly 1 percent for rest of authors in top 25 category.

Table 4.2: Authorship Pattern (JNU)

Field: Author	Record Count	% of 6,344
Prasad, R.	126	1.986%
Ramaswamy, R.	124	1.954%
Baquer, N.Z.	118	1.860%
Bhattacharya, A.	117	1.844%
Puri, S.	112	1.765%
Bohidar, H.B.	112	1.765%
Subramanian, V.	94	1.481%
Ramanathan, A.L.	94	1.481%
Behari, J.	92	1.450%
Sopory, S.K.	86	1.355%
Mohanty, P.	84	1.324%
Bhattacharya, S.	82	1.292%
Madhubala, R.	76	1.197%
Datta, K.	74	1.166%
Kale, R.K.	74	1.166%
Ghosh, S.	73	1.150%
Ghosh, R.	73	1.150%
Bhatnagar, R.	70	1.103%
Saxena, R.K.	70	1.103%
Mallick, B.N.	69	1.087%
Bamezai, R.N.K.	66	1.040%
Rao, A.R.	65	1.024%
Kesavan, P.C.	62	0.977%
Patnaik, S.	58	0.914%
Lobiyal, D.K.	55	0.866%

The authorship pattern of JNU is widely dispersed among several author and unlike University of Delhi there was very less concentration. Most of the top 25 authors have contributed close to 1 percent of total research output.

Table 5: Publication Year (DU)

Field: Publication Years	Record Count	% of 17,128
2014	425	2.481%
2013	1208	7.052%
2012	1265	7.385%
2011	1195	6.976%
2010	1050	6.130%
2009	847	4.945%
2008	778	4.542%
2007	688	4.016%
2006	591	3.450%
2005	548	3.199%

2004	469	2.738%
2003	489	2.854%
2002	371	2.166%
2001	353	2.060%
2000	272	1.588%
1999	317	1.850%
1998	330	1.926%
1997	287	1.675%
1996	326	1.903%
1995	184	1.074%
1994	176	1.027%
1993	188	1.097%
1992	168	0.980%
1991	190	1.109%

If we look at the yearly distribution of research papers, we are able to see an increasing trend in the number of papers. This establishes more research efforts and better facilities for undertaking research in the university.

Table 5.1: Publication Year (JMI)

Field: Publication Year	**Record Count**	**% of 998**
2014	54	5.410%
2013	134	13.426%
2012	147	14.729%
2011	138	13.827%
2010	118	11.823%
2009	79	7.915%
2008	68	6.813%
2007	72	7.214%
2006	62	6.212%
2005	45	4.509%
2004	20	2.004%
2003	10	1.002%
2002	6	0.601%
2001	2	0.200%
2000	10	1.002%
1999	4	0.400%
1998	8	0.801%
1997	5	0.501%
1996	3	0.300%
1995	3	0.300%
1994	1	0.100%
1991	2	0.200%

At JMI also the number of research publication has gone up steadily every year.

Table 5.2: Publication Year (JNU)

Field: Publication Year	Record Count	% of 6,344
2014	190	2.994%
2013	491	7.739%
2012	525	8.275%
2011	479	7.550%
2010	390	6.147%
2009	358	5.643%
2008	345	5.438%
2007	242	3.814%
2006	226	3.562%
2005	258	4.066%
2004	240	3.783%
2003	195	3.073%
2002	170	2.679%
2001	178	2.805%
2000	159	2.506%
1999	169	2.663%
1998	185	2.916%
1997	138	2.175%
1996	158	2.490%
1995	87	1.371%
1994	92	1.450%
1993	123	1.938%
1992	86	1.355%
1991	98	1.544%
1990	92	1.450%

The largest number of articles have been published in year 2012 constituting more than 5 percent of total research. However the distribution is quite evenly spread with slight increase every year.

Table 6: Countries/Territories (DU)

Field: Countries/Territories	Record Count	% of 17,128
India	16599	96.911%
United States	1834	10.707%
United Kingdom	993	5.797%
Germany	888	5.184%
France	822	4.799%
South Korea	771	4.501%
Russian Federation	749	4.372%
China	736	4.297%
Brazil	723	4.221%
Mexico	713	4.162%
Colombia	661	3.859%
Czech Republic	570	3.327%

Switzerland	539	3.146%
Italy	450	2.627%
Spain	439	2.563%
Canada	428	2.498%
Ecuador	394	2.300%
Ireland	382	2.230%
Poland	381	2.224%
Netherlands	380	2.218%
Sweden	379	2.212%
Argentina	359	2.095%
Finland	353	2.060%
Belgium	344	2.008%
Taiwan	330	1.926%

If we look at the collaborating authors in terms of their affiliating countries, a large number of papers have been written in collaboration with several countries. However USA, UK, Germany, France and South Korea, Russia, China and Brazil are main collaborating countries with US being at top contributing nearly 10 percent.

Table 6.1: Countries/Territories (JMI)

Field: Countries/Territories	Record Count	% of 998
India	975	97.695%
Saudi Arabia	97	9.719%
United States	27	2.705%
Egypt	27	2.705%
Japan	23	2.304%
Taiwan	15	1.503%
Germany	14	1.402%
Iraq	13	1.302%
Australia	10	1.002%
United Kingdom	10	1.002%
Malaysia	9	0.901%
Poland	8	0.801%
Oman	7	0.701%
Portugal	6	0.601%
Iran	6	0.601%
United Arab Emirates	5	0.501%
Italy	5	0.501%
Canada	5	0.501%
France	5	0.501%
South Korea	4	0.400%
Thailand	4	0.400%
Netherlands	3	0.300%
Belgium	3	0.300%
South Africa	3	0.300%
Switzerland	3	0.300%

If we look at the collaborating authors in terms of their affiliating countries for Jamia Millia Islamia, Saudi Arabia occupies topmost position constituting nearly 10 percent of total. It is followed by US, Egypt, Taiwan, Germany and Iraq.

Table 6.2: Countries/Territories (JNU)

Field: Countries/Territories	Record Count	% of 6,344
India	6089	95.980%
United States	533	8.401%
Germany	174	2.742%
United Kingdom	115	1.812%
France	83	1.308%
Japan	69	1.087%
Canada	56	0.882%
Australia	35	0.551%
Italy	31	0.488%
Sweden	31	0.488%
Netherlands	30	0.472%
Switzerland	27	0.425%
China	26	0.409%
Belgium	25	0.394%
South Korea	25	0.394%
Russian Federation	18	0.283%
Spain	17	0.267%
Israel	16	0.252%
Mexico	14	0.220%
Singapore	13	0.204%
Poland	12	0.189%
Malaysia	11	0.173%
Finland	10	0.157%
Saudi Arabia	10	0.157%
Austria	10	0.157%

With regard to JNU, highest collaboration has taken place with US (8.33%) followed by Germany, UK, France, Japan and Canada

Table 7: Institutions (DU)

Institutions	Record Count	% of 17,128
University of Delhi	17128	100.00%
Punjab University	799	4.664%
Tata Institute of Fundamental Research	755	4.407%
Fermi National Accelerator Laboratory	723	4.221%
University of Rochester	712	4.156%
Northeastern University	710	4.145%
University of Notre Dame	699	4.081%
Institute fiziki vysokikh	694	4.051%
University of California, Riverside	693	4.046%
University of Illinois at Chicago	692	4.040%

Florida State University	687	4.010%
Brown University	687	4.010%
Rice University	684	3.993%
Northwestern University	684	3.993%
Centro Brasileiro de Pesquisas Fisicas	679	3.964%
Centro de Investigacion y de Estudios Avanzados	677	3.952%
CEA Saclay	674	3.935%
University of Nebraska - Lincoln	666	3.888%
Universidade do Estado do Rio de Janeiro	665	3.882%
Universidad de Los Andes	659	3.847%
Moskovskij Gosudarstvennyj Universitet	644	3.759%
Alikhanov Institute for Theoretical and Experimental Physics	642	3.748%
Boston University	635	3.707%
University of Kansas Lawrence	629	3.672%
Korea University	613	3.578%

With regard to research collaboration in terms of Institutions Punjab University, Tata Institute of Fundamental Research, Fermi National Accelerator Laboratory, University of Rochester, Northeastern University, University of Notre Dame Institute fiziki vysokikh University of California, Riversi University of Illinois at Chicago Florida State University and university Brown University are topmost intuitions contributing nearly 4 percent of over all research output.

Table 7.1: Institutions (JMI)

Institutions	Record Count	% of 998
Jamia Millia Islamia	998	100.00%
King Abdulaziz University	54	5.410%
University of Delhi	54	5.410%
Indian Institute of Technology, Delhi	53	5.310%
National Physical Laboratory India	42	4.208%
National Research Center, Cairo	26	2.605%
Indian Institute of Technology Roorkee	24	2.404%
Aligarh Muslim University	19	1.903%
Jawaharlal Nehru University	16	1.603%
Guru Gobind Singh Indraprastha University	15	1.503%
King Saud University College of Science	14	1.402%
Indian Agricultural Research Institute	14	1.402%
King Saud University	13	1.302%
Jamia Hamdard University	11	1.102%
Inter University Accelerator Centre India	9	0.901%
Jadavpur University	9	0.901%
Solid State Physics Laboratory India	8	0.801%
Krishna Institute of Engineering and Technology	8	0.801%
Central Glass and Ceramic Research Institute India	7	0.701%
Kyushu Institute of Technology	7	0.701%
Chaudhary Devi Lal University, Sirsa	7	0.701%

Defence Research and Development Organisation India	6	0.601%
International Centre for Genetic Engineering and Biotechnology	6	0.601%
Ministry of Science and Technology	5	0.501%
Prochrome India	5	0.501%

With regard to research collaboration in terms of Institutions for Jamia Millia Islamia, King Abdulaziz University of Delhi and Indian Institute of Technology, Delhi are top most intuitions contributing nearly 5 percent of over all research output.

Table 7.2: Institutions (JNU)

Institutions	Record Count	% of 6,344
Jawaharlal Nehru University	6344	100.00%
University of Delhi	171	2.695%
Indian Institute of Technology, Delhi	117	1.844%
All India Institute of Medical Sciences	90	1.418%
International Centre for Genetic Engineering and Biotechnology, New Delhi	58	0.914%
International Centre for Genetic Engineering and Biotechnology	54	0.851%
Govind Ballabh Pant Institute of Himalayan Environment and Development	44	0.693%
Inter University Accelerator Centre India	42	0.662%
National Institute of Immunology India	40	0.630%
Tata Institute of Fundamental Research	39	0.614%
Annamalai University	36	0.567%
Indian Agricultural Research Institute	35	0.551%
Universite Paris-Sud XI	34	0.535%
Johannes Gutenberg Universitat Mainz	28	0.441%
Aligarh Muslim University	28	0.441%
University Colorado Cancer Center	27	0.425%
National Physical Laboratory India	27	0.425%
Indian Institute of Science	27	0.425%
Bhabha Atomic Research Centre	27	0.425%
Banaras Hindu University	26	0.409%
University of Colorado Health Sciences Center	24	0.378%
Jawaharlal Nehru Centre for Advanced Scientific Research	23	0.362%
Central University of Gujarat	21	0.331%
Guru Nanak Dev University India	21	0.331%
Indian Institute of Technology, Kanpur	21	0.331%

With regard to research collaboration in terms of Institutions in respect of JNU, University of Delhi Indian Institute of Technology, Delhi and AIIMS are top most intuitions contributing nearly 1 percent of over all research output.

Table 8: Source Titles (DU)

Source Titles	Record Count	% of 17,128
Physical Review Letters	293	1.710%
Physical Review D Particles Fields Gravitation and Cosmology	219	1.278%
Physics Letters Section B Nuclear Elementary Particle and High Energy Physics	204	1.191%
Physical Review D	195	1.138%
Journal of Applied Physics	170	0.992%
Phytochemistry	165	0.963%
Physical Review	163	0.951%
Current Science	151	0.881%
Journal of the Indian Chemical Society	139	0.811%
Indian Journal of Experimental Biology	135	0.788%
Proceedings of the Indian Academy of Sciences Section A	134	0.782%
Spectrochimica Acta Part A Molecular and Biomolecular Spectroscopy	123	0.718%
Indian Journal of Chemistry Section B Organic and Medicinal Chemistry	120	0.700%
Economic and Political Weekly	107	0.624%
Tetrahedron Letters	103	0.601%
Indian Journal of Chemistry Section A Inorganic Physical Theoretical and Analytical Chemistry	91	0.531%
Journal of High Energy Physics	90	0.525%
Tetrahedron	89	0.519%
Journal of Physics D Applied Physics	88	0.513%
Nature	87	0.507%
Transition Metal Chemistry	83	0.484%
Proceedings of SPIE the International Society for Optical Engineering	80	0.467%
Pramana Journal of Physics	78	0.455%
Physics Letters A	75	0.437%
Thin Solid Films	71	0.414%

With regard to Source Titles of research output of University of Delhi, Physical Review Letters, Physical Review D, Particles Fields Gravitation and Cosmology, Physics Letters Section B, Nuclear Elementary Particle and High Energy Physics Physical Review D are top most titles constituting nearly 1 percent of source titles.

Table 8.1: Source Titles (JMI)

Source Titles	Record Count	% of 998
Proceedings of SPIE the International Society for Optical Engineering	12	1.202%
Colloids and Surfaces B Biointerfaces	12	1.202%
Economic and Political Weekly	9	0.901%
Journal of Interdisciplinary Mathematics	9	0.901%
Journal of the Indian Chemical Society	8	0.801%
Colloids and Surfaces A Physicochemical and Engineering Aspects	8	0.801%

Lecture Notes in Computer Science Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics	8	0.801%
Journal of Applied Polymer Science	8	0.801%
Journal of Molecular Liquids	8	0.801%
Journal of Commonwealth Literature	7	0.701%
Arabian Journal of Chemistry	7	0.701%
Physica C Superconductivity and Its Applications	6	0.601%
Colloid and Polymer Science	6	0.601%
Physical Review D Particles Fields Gravitation and Cosmology	6	0.601%
Aip Conference Proceedings	6	0.601%
Journal of Coordination Chemistry	6	0.601%
Journal of Dispersion Science and Technology	6	0.601%
International Journal of Nanoparticles	6	0.601%
Journal of Applied Physics	6	0.601%
Acta Physico Chimica Sinica	5	0.501%
International Journal of Thermo physics	5	0.501%
World Academy of Science Engineering and Technology	5	0.501%
Combinatorial Chemistry and High Throughput Screening	5	0.501%
Communications in Computer and Information Science	5	0.501%
Physics and Chemistry of Liquids	5	0.501%

With regard to Source Titles of research output of Jamia Millia Islamia , Proceedings of SPIE the International Society for Optical Engineering, Colloids and Surfaces B Biointerface, Economic and Political Weekly and Journal of Interdisciplinary Mathematics are top most titles constituting nearly 1 percent of source titles.

Table 8.2: Source Titles (JNU)

Source Titles	**Record Count**	**% of 6,344**
Economic and Political Weekly	233	3.672%
International Studies	133	2.096%
Indian Journal of Experimental Biology	99	1.560%
Physical Review E Statistical Nonlinear and Soft Matter Physics	97	1.529%
Current Science	89	1.402%
Biochemical and Biophysical Research Communications	82	1.292%
Plos One	63	0.993%
Journal of Biosciences	50	0.788%
Journal of Biological Chemistry	49	0.772%
Biochemistry International	47	0.740%
Indian Journal of Labour Economics	46	0.725%
Journal of Chemical Physics	45	0.709%
China Report	45	0.709%
Indian Journal of Biochemistry and Biophysics	44	0.693%
Studies in History	40	0.630%

Lecture Notes in Computer Science Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics	40	0.630%
Environmental Monitoring and Assessment	39	0.614%
South Asian Survey	37	0.583%
Physical Review B Condensed Matter and Materials Physics	37	0.583%
Pramana Journal of Physics	32	0.504%
International Journal of Radiation Biology	32	0.504%
FEMS Microbiology Letters	31	0.488%
Cancer Letters	29	0.457%
Applied Physics Letters	28	0.441%
Journal of Physics Condensed Matter	27	0.425%

With regard to Source Titles of research output of JNU, Economic and Political Weekly , an Indian journal is top most source title and International Studies, Indian Journal of Experimental Biology, Physical Review E Statistical Nonlinear and Soft Matter Physics, Current Science, Biochemical and Biophysical Research Communications are other top titles constituting nearly 3 to 1 percent of source titles of total research output.

Citation Analysis

Citation analysis is done to know how many times a research paper has been cited by subsequent researchers. Subsequent citations establish the usefulness and importance of a research paper and add value to it. The citation analysis for all three universities is given below.

University of Delhi:

Total Number of Papers Published: 17,128

Total Number of Citations since 1998: 138846

Average Citation per paper: 8.10

Jamia Milia Islamia

Total Number of Papers Published: **1008**

Total Number of Citations since 1998: 5808

Average Citation per paper: 5.76

Jawahal Lal Nehru University:

Total Number of Papers Published: 6,344

Total Number of Citations since 1998: 54747

Average Citation per paper: 8.62

Findings

- Yearly average of papers published by University of Delhi is 217 for last 79 years for which data is available.
- Yearly average of papers published by Jamia Millia is 25 for last 41 years for which data is available.
- Yearly average of papers published by JNU is 140 for last 45 years for which data is available.
- Thus the average publishing rate of Delhi University is Highest.

- Highest number of papers have been published by University of Delhi in the area of Physics and Astronomy followed by Chemistry and Biotechnology, Genetics and Molecular Biology. The least number of papers published are in the area of Dentistry.
- The subject wise distribution of JMI reveals that highest number of papers have been published are in the area of Chemistry followed by Engineering, Physics and Astronomy.
- The highest number of papers by JNU have been published in the interdisciplinary area of Chemistry , Engineering , Physics and Astronomy followed by Social Sciences and Physics and Astronomy. It clearly discernible from data JNU is quite strong in the area of social science research in comparison to University of Delhi.
- 23 percent publications of University of Delhi are in the form of articles which is most established format for research output and followed by conference papers.
- The 73.5 percent publications of JMI are in the form of articles whereas 16.6 percent research is in the form of conference papers.
- 78. 8 percent of JNU research has been published in the form of articles whereas 6.99 percent in the form of conference papers.
- The authorship pattern of University of Delhi reveals that topmost author has contributed slightly over four percent to the research profile of University of Delhi whereas 25 authors have contributed more than 3 percent.
- The authorship pattern of Jamia Millia Islamia looks less concentrated though top most author has contributed more than 7 percent of total research followed by two authors contributing 6 percent or so each. The percentage get down to nearly 1 percent for rest of authors in top 25 category.
- The authorship pattern of JNU is widely dispersed among several author and unlike University of Delhi there was very less concentration. Most of the top 25 authors have contributed close to 1 percent of total research output.
- The yearly distribution of research papers of Delhi Unibversity reveals an increasing trend in the number of papers. This establishes more research efforts and better facilities for undertaking research in the university.
- At JMI the number of research publication has gone up steadily every year.
- The largest number of articles by JNU have been published in year 2012 constituting more than 5 percent of total research. However the distribution is quite evenly spread with slight increase every year.
- At Delhi University USA, UK, Germany, France and South Korea, Russia, China and Brazil are main collaborating countries with US being at top contributing nearly 10 percent.
- The collaborating authorship pattern in terms of their affiliating countries for Jamia Millia Islamia, reveals that Saudi Arabia occupies topmost position constituting nearly 10 percent of total. It is followed by US, Egypt, Taiwan, Germany and Iraq.
- With regard to JNU, highest collaboration has taken place with US (8.33%) followed by Germany, UK, France, Japan and Canada
- With regard to Source Titles of research output of University of Delhi, Physical Review Letters, Physical Review D , Particles Fields Gravitation and Cosmology, Physics Letters Section B, Nuclear Elementary Particle and High Energy Physics Physical Review D are top most titles constituting nearly 1 percent of source titles.
- With regard to Source Titles of research output of Jamia Millia Islamia , Proceedings of SPIE the International Society for Optical Engineering, Colloids and Surfaces B Biointerface, Economic and Political Weekl and Journal of Interdisciplinary Mathematics are top most titles constituting nearly 1 percent of source titles.
- With regard to Source Titles of research output of JNU , Economic and Political Weekly , an Indian journal is top most source title and International Studies, Indian Journal of

Experimental Biology, Physical Review E Statistical Nonlinear and Soft Matter Physics, Current Science, Biochemical and Biophysical Research Communications are other top titles constituting nearly 3 to 1 percent of source titles of total research output.
- Topmost Indian journal is Economic and Political Weekly contributing to social science research impressively.

Citation Analysis

The highest number of paper have been published by University of Delhi both in absolute term and also on average basis but the highest number of citation have been received by JNU ie 8.62 closely followed by University of Delhi which is 8.10.

Conclusion

Thus it can be argued that there exists sizeable body of research among three Universities located in the capital city of India covering a wide variety of subjects and disciplines. The research papers have widely been acknowledge in terms of their contribution and cited by subsequent researchers quite substantially establishing the credibility and importance of research.

References and Bibliography

Adarsh Bala & Gupta, B.M. (2009). Growth and impact of research output of government medical college & hospital, Chandigarh: a case study. *"Annals of Library and Information Studies, 56"*, 86-94.

Bhat, Veena R. & Sampath Kumar, B.T. (2008). Use of web based sources in scholarly electronic journals in the field of library and information science: a citation analysis. *"Annals of Library and Information Studies, 55"*, 145-152.

Biradar, B.S. (2006). Indian journal of environmental protection: a study of citation pattern. *"Annals of Library and Information Studies, 53"*, 109-113.

Garfield, Eugene (1988)."*Journal citation studies 47"*. Which oceanography journals make the biggest waves? IN SCI Journal Citation Reports. Philadelphia, Institute for Scientific Information Inc. 1988.

Garg, K.C. and Kumar, S. (2010). An analysis of the citation pattern of Indian science journals indexed by SCIE. *"Annals of Library and Information Studies, 57"*, 365-372

Garg, K.C., Dutt, B. & Suresh Kumar (2006). Scientometric profile of Indian science as seen through Science Citation Index. *"Annals of Library and Information Studies, 53"*, 114-125.

Gopalakrrishnan, S., Ambuja, R. & Seetharama, S. (2002). A webometric study on Library and Information Science literature. *"Annals of Library and Information Studies, 49 (1)"*, 29-35.

Kumbar, Mallinath, Gupta, B.M. & Dhawan, S.M. (2008). Growth and impact of research output of University of Mysore, 1996-2006: A case study. *"Annals of Library and Information Studies, 55",* 185-195.

Peters, H.P.F., Hartmann, D. & Raan , A.F.J. Van. *"Monitoring advances in chemical engineering, IN Informetrics 87/88*", L.Egghe and R.Rousseau (eds.), Amsterdam, Elsevier Science Publishers, 1988.

Pillai, K.G.Sudhier (2007). Authorship pattern in physics literature: an informetric study on citations in doctoral theses of the Indian Institute of Science. *"Annals of Library and Information Studies, 54"*, 90-94.

Raut, T.K., Sahu, S.B. & Ganguly, S. (2008). Strategic Management Journal: a citations study. *Annals of Library and Information Studies, 55"*, 69-75.

Rousseau, Ronald(1988). *"Citation distribution of pure mathematics journals, in Informetrics 87 / 88"*. L.Egghe and R.Rousseau (eds.) Amsterdam, Elsevier, Science Publishers, 1988.

Spies, Brian R. (1991). *"The effectiveness of journals in exploration, geophysics 1991 vol. 56 no. 6 (June)"*, 844-858.

Surendra Kumar & S.Kumar (2008). Citation analysis of Journal of Oilseeds Research 1993-2004. *"Annals of Library and Information Studies, 55"*, 35-44.

Uzun, A., A. Menard, and M.E. Ozel(1993)."*Citation status of Turkish physics foreign journals: A global analysis". Scientometrics,1993 vol. 28 no. 1*, 79-87.

Yogendra Singh & Gupta, B.M. (2005). Research contributions and impact of research of IIT, Roorkee, 1993 to 2001. *"Annals of Library and Information Studies, 52 (1)"*, 8-14.

International research collaboration of Sri Lanka in the last 02 decades (1994 – 2013) based on the SCOPUS database

Thuraiyappah Pratheepan* and W.A. Weerasooriya**

*Library, Uva Wellassa University, Badulla. Sri Lanka
pratheepan12345@gmail.com

**Department of Library & Information Science, University of Keleniya, Sri Lanka

Abstract

This paper aims to investigate & analyze Sri Lanka's International Research Collaboration from 1994 – 2013 using scientometric methods. The SCOPUS electronic database was used to compile scientific collaboration of Sri Lanka with other countries for the period, 1994 – 2013.Sri Lanka has engaged in International Research Collaboration with 159 countries. Sri Lanka has therefore engaged with 80% of the countries in Research Collaboration. This is a very healthy sign for a developing country like Sri Lanka to reach the quality of research. There is a steady increase in collaborative research of Sri Lanka with other countries during the study period. Sri Lanka's partners are the leading economies in the world. United Kingdom is the country which has the most predominant collaboration to share with Sri Lanka with a count of 1205(15.77%), followed by USA – 1093 (14.30%), Australia – 805 (10.53%) and in the fourth and fifth position Japan and India have the remarkable collaboration to share with Sri Lanka with the count of 668 (8.74%) and 503 (6.58%) respectively. In addition, Statistics reveal that Sri Lankan researchers have collaborated with the world class universities, R&D institutions significantly. This is a very healthy sign for a developing country like Sri Lanka in the field of research.

Keywords: International partnership, Scientometrics, Scientific Collaboration, Sri Lanka.

Introduction

The government of Sri Lanka sees international research collaboration as imperative for the nation in several ways. Because collaborative research programmes help to stretch Sri Lanka's limited budget. Research collaboration with other countries helps to maintain world – class scientific standards. International research collaboration promotes the exchange of information and ideas. In general Sri Lanka has a tradition of collaborative projects. This paper aims to investigate & analyze Sri Lanka's international research collaboration in SCOPUS database from 1994 – 2013 using scientometric methods. International research collaboration appears to be very essential for developing countries like Sri Lanka. Collaboration must include European and US researchers in Asian and South American laboratories (Adams, 2013, p.560). Recently great importance has been shown for international research collaboration in Sri Lanka. Choosing the key performance indicator *(KPI)* of international research collaboration by the Higher Education sector confirms this. On this basis evaluating Sri Lanka's international research partnership through scientific publications is significant because it will give awareness of further and future development among scientists, R&D institutions, Higher Education authorities etc. to achieve the common goal of producing new scientific knowledge through effective collaboration with other countries.

It is very clear that Impact of research collaboration with other nations has been influencing the quality of research in many ways. Writing a research paper in collaboration with other nations, departments or/and institutions can be very productive. Since 1901, the percentage of Nobel prizes awarded to two or three individuals for one project has increased over the years going from 14.8% in the 1900s to >60% in the 1970s,1980, and 1990s gives an indication of the value and frequency of collaborative work (Hafernik, Messerschmitt & Vandrick , 1997, p.31). Impact is typically greater when research groups collaborate, and the benefit strengthens when co – authorship is international (Schmouch & Schubert, 2008,p.363). International research collaboration, which, it is argued, is characteristic of rapidly changing research systems, is regarded by many as an indicator of high-quality research (Kim, 2006, p.231). International collaboration encourages best practice through the sharing of ideas and facilitates evidence-based practice (Freshwater, Sherwood, Drury, 2006, p.302). Bookstein, Moed and Yitzahki (2006) measured the strength of international collaboration, as reflected by the co-authorship of research papers. Collaboration helps individuals into doing research. Exploring issues of mutual interest by doing collaborative work is not only personally and professionally enriching but also benefit the field as a whole ((Hafernik, Messerschmitt & Vandrick , 1997). Further, Anuradha and Urs (2007) analyzed patterns in research collaboration in India. Collaborative Program with foreign countries is Sri Lanka's most longstanding national arrangement for universities-R&D institutes - government research.

Over the past two decades international collaboration has grown to become the dominant model for R&D activities in Sri Lanka.Up to very recently, none of the Sri Lankan universities or research institutions had the entire necessary infrastructure for internal research. Making Sri Lanka a hub of higher education in the international arena with the least possible expense appears to be the objective of the Ministry of Higher Education and the University Grants Commission. Under the new vision of making Sri Lanka the most effective centre of quality education & research in South Asia the Ministry of Higher Education has set several goals in order to achieve this objective. In this regard international research collaboration is vital for Sri Lanka in order to achieve research excellence.

Statement of research problem

No study has been conducted on collaborative research of Sri Lanka with other nations. It was decided to undertake the study on International Research Collaboration of Sri Lanka during 1994-2013 for a period of 20 years based on the SCOPUS database.

Methodology

The SCOPUS electronic database was used to compile scientific collaboration of Sri Lanka with other countries for the period, 1994 – 2013 using scientometric methods. The use of advanced search in which we used the 'field tag for country' search for Sri Lanka - Sri Lanka or Ceylon because until 1972 the country was officially known as Ceylon, few of the articles published were in the name Ceylon. The 9807 results retrieved are analyzed to map Sri Lanka's international research collaboration. Out of 159 collaborators Sri Lanka's top 10 leading international research collaborators in the last 02 decades were selected for detailed investigation.

Results & Discussion

The total number of research output of Sri Lanka in the SCOPUS database from 1994 to 2013 consists of 9807 papers. This includes all types of documents such as journal articles, conference proceedings, abstracts etc. The Sri Lankan research outputs have been produced in collaboration with 159 countries in the world. There is a steady increase in collaborative research of Sri Lanka with other countries during the years 1994-2013. Table 1 highlights the level of Sri Lanka's International collaborations. It shows numbers of publications with overseas addresses. The collaborations are shown for four time periods: 1994-1998, 1999–2003, 2004 - 2008 and 2009 – 2013. Table 1 shows top 10 collaborators of Sri Lanka.

Table 1.

Sri Lanka's top 10 leading research partners in the last 02 decades

	Collaborative papers				
Leading partners	**1994-1998**	**1999-2003**	**2004-2008**	**2009-2013**	**Total**
United Kingdom	124	188	410	483	**1205**
United States	95	145	366	487	**1093**
Australia	46	47	216	496	**805**
Japan	35	94	244	295	**668**
India	11	38	154	300	**503**
Canada	26	31	108	141	**306**
Sweden	18	29	90	95	**232**
Germany	12	37	81	99	**229**
Thailand	11	20	60	123	**214**
Netherlands	15	32	63	92	**202**

Sri Lanka's partners are the leading economies in the world. United Kingdom is the country which has the most predominant collaboration to share with Sri Lanka with a count of 1205(15.77%), followed by USA – 1093 (14.30%), Australia – 805 (10.53%) and in the fourth and fifth position Japan and India have the remarkable collaboration to share with Sri Lanka with the count of 668 (8.74%) and 503 (6.58%) respectively. The scientific collaboration of Sri Lanka with UK has risen from 124 in period one (1994-1998) to 483 in period four (2009-2013) illustrates that the collaborative work of Sri Lanka has received sustained attention for technological innovation. As the predominant collaborator with a count of 1205 UK has published its collaborative work with Sri Lanka in 160 different sources. Of these 939 were full articles, 97 Review, 89 conference papers, 80 other communications.

As shown in Table 2 amongst the varied type of documents, journal articles found to be the chief carrier of collaborative research communications and the other forms of scientific communication are comparatively less.

Table 2.

Document type selected for collaborative publication from leading partners of Sri Lanka

Document Type	UK	USA	Australia	Japan	India	Canada	Sweden	Germany	Thailand	Netherland
Journal Article	77.93%	80.24%	74.41%	79.49%	80.12%	76.14%	74.14%	82.53%	76.64%	76.73%
Conference Paper	7.39%	8.97%	13.17%	15.27%	5.77%	11.11%	18.10%	7.42%	10.75%	8.42%
Review	8.05%	6.59%	8.45%	3.59%	8.95%	6.86%	3.88%	5.68%	6.54%	6.93%
Others	6.64%	4.21%	3.98%	1.65%	5.17%	5.88%	3.88%	4.37%	6.07%	7.92%
Total	**1205**	**1093**	**805**	**668**	**503**	**306**	**232**	**229**	**214**	**202**

Table 3 shows the first three preferences of sources that have been selected by researchers from leading countries for collaborative publication with Sri Lanka. When selecting the sources where their research findings are to be published; the standard of the journal has to be kept in mind. *"Lancet*", according to our findings is the Journal that has published the most number of research papers of 07 countries.

Table 3.

First three preferences of sources from leading countries for collaborative publication with Sri Lanka.

Countries	1st Preference	2nd Preference	3rd Preference	No. of Sources used
UK	Lancet **(32)**	$ Tran.of the R. S T.M&H Hygiene **(17)**	Plos One **(17)**	160
USA	Lancet **(17)**	$$ Phy. Rev. Mat.& Mat. Phys. **(14)**	Zootaxa **(13)**	157
Australia	Lancet **(22)**	Clinical Toxicology **(22)**	BMC Public Health **(10)**	149
Japan	Lancet **(11)**	Parasitology International **(10)**	Soil Science and Plant Nutrition **(9)**	152
India	# Act. Crst.Sec E Stru. Rep **(75)**	Lancet **(14)**	Malaria Journal **(9)**	118
Canada	Lancet **(13)**	$$$ W.W.Hosp. and Touri. Themes **(8)**	Science **(5)**	147
Sweden	## J .of Atmos.Phys.**(9)**	Lancet **(6)**	Solid State Ionics **(6**]	145
Germany	Lancet **(8)**	Chemosphere **(5)**	* Envi. Geochemis. and Health **(5)**	136
Thailand	Lancet **(11)**	Energy Conversion and Management **(5)**	Aquaculture **(5)**	127
Netherland	### Agri. Water Mgt. **(12)**	$$$$ Asian Aus Jour. of ANS **(8)**	Lancet **(8)**	114

\# *Acta Crystallographica Section E Structure Reports Online*
\## *Journal of Atmospheric Physics*
\### *Agricultural Water Management*
$ *Transactions of the Royal Society of Tropical Medicine and Hygiene*
$$ *Physical Review B Condensed Matter and Materials Physics*
$$$ *Worldwide Hospitality and Tourism Themes*
$$$$ *Asian Australasian Journal of Animal Sciences*
* *Environmental Geochemistry and Health (5)*

One might expect considerable difference in discipline - collaboration patterns. Through this collaboration, medicine proved to be Sri Lanka's key research area of highest representation for collaborative research in the SCOPUS database, with 523 papers followed by Agriculture and Environmental Sciences with the count of 284 and 185 respectively.

There are 34 institutions that have contributed for collaborative publication with other countries in the last two decades from Sri Lanka. Among the institutions of Sri Lanka involved in international research partnership, the University of Peradeniya, University of Colombo, and International Water Management Institute are found to be the predominant institutes. Table 4 lists the Sri Lankan institutions that have contributed substantially to the collaborative literature with top 10 leading countries.

Table.4

Key collaborating institutions of Sri Lanka with top 10 collaborating countries

	Sri Lanka's Institutes	Record
Leading collaborators with USA	University of Peradeniya	269
	University of Colombo	177
	IWMI	103
	Institute of Fundamental Studies	92
	University of Ruhuna	65
Leading collaborators with Australia	University of Colombo	203
	University of Peradeniya	196
	University of Kelaniya	57
	IWMI	55
	University of Sri Jayewardenepura	34
Leading collaborators with UK	University of Colombo	299
	University of Peradeniya	259
	University of Kelaniya	110
	University of Sri Jayewardenepura	73
	IWMI	55
Leading collaborators with Japan	University of Peradeniya	263
	Institute of Fundamental Studies	79
	University of Ruhuna	54
	University of Kelaniya	45
	University of Moratuwa	41
Leading collaborators with India	University of Ruhuna	67
	IWMI	65
	University of Peradeniya	63
	University of Colombo	61
	Eastern University	28
Leading collaborators with Canada	University of Peradeniya	91
	University of Colombo	39
	University of Kelaniya	33
	University of Moratuwa	27
	IWMI	15
Leading collaborators with Sweden	University of Colombo	87
	University of Peradeniya	68
	Institute of Fundamental Studies	14
	University of Ruhuna	11
	University of Sri Jayewardenepura	10
Leading collaborators with Germany	University of Peradeniya	66
	Institute of Fundamental Studies	42
	University of Colombo	38
	IWMI	20
	University of Kelaniya	11
Leading collaborators with Thailand	University of Peradeniya	34
	University of Kelaniya	26
	University of Colombo	20
	University of Moratuwa	15
	IWMI	10
Leading collaborators with Netherland	IWMI	57
	University of Peradeniya	33
	University of Ruhuna	21
	University of Kelaniya	19
	University of Colombo	18

Table. 5

Collaboration of leading institutions with Sri Lanka from top 10 leading countries

	Foreign Institutes Vs Sri Lanka	Record
USA's Leading collaborators	Yale University	39
	Harvard University	32
	University of Houston	31
	Cornell University	28
	UC Davis	25
Australia's Leading collaborators	University of New South Wales	74
	University of Melbourne	72
	University of Queensland	72
	University of Sydney	58
	Monash University	54
UK's Leading collaborators	King's College London	99
	University of Oxford	85
	University of Edinburgh	63
	University of Manchester	52
	Imperial College London	51
Japan's Leading collaborators	University of Tokyo	56
	Tokyo Institute of Technology	34
	Obihiro University	33
	Saitama University	24
	Saga University	23
India's Leading collaborators	Madurai Kamaraj University	49
	Cochin University of Science and Technology	27
	Christian Medical College, Vellore	19
	All India Institute of Medical Sciences	17
	Indian Institute of Science	17
Canada's Leading collaborators	The University of British Columbia	40
	University of Calgary	25
	University of Manitoba	24
	University of Toronto	23
	University of Guelph	19
Sweden's Leading collaborators	Uppsala Universitet	36
	Chalmers Tekniska Högskola	31
	The Royal Institute of Technology KTH	26
	Stockholms universitet	20
	Göteborgs Universitet	19
Germany's Leading collaborators	Universität Erlangen-Nürnberg	28
	Medizinische Fakultät der LMU München	14
	Ludwig-Maximilians-Universität München	12
	Universität Bremen	11
	Universität Bayreuth	9
Thailand's Leading collaborators	Asian Institute of Technology	53
	Chulalongkorn University	31
	Mahidol University	27
	Chiang Mai University	12
	Thammasat University	12
Netherland's Leading collaborators	Wageningen University	46
	Delft University of Technology	19
	Utrecht University	15
	University of Twente	10
	Vrije Universiteit Amsterdam	9

Table 5 provides data on papers published by top 5 international institutions from leading collaborative countries. Sri Lankan researchers have collaborated with 120 R&D institutes including universities from 159 countries during the last two decades. When International Research Collaboration is undertaken, the importance of judging the international standard of R&D institution and university is very relevant. Thankfully the top 10 leading research partners who have been selected by the Sri Lankan scientists are world's strongest countries in R&D activities. In addition, Statistics reveal that Sri Lankan researchers have collaborated with the world class universities, R&D institutions significantly. This is a very healthy sign for a

Table 6.

Sri Lanka's top 3 prolific collaborators with each leading partner countries and their affiliating institute

No	Authors with UK	Output	Affiliation	Field
1	Sheriff, M.H.R.	26	University of Colombo	Medicine
2	Malavige, G.N.	26	University of Sri Jayawardanepura	Medicine
3	Ekanayake, J.B.	25	University of Peradeniya	Engineering
	Authors with USA			
1	Gunaratne, G.H.	31	Institute of Fundamental Studies	Physics
2	Meegaskumbura, M.	16	Wildlife Heritage Trust	Envirmt.Sc.
3	Singhakumara, B.M.P.	13	University of Sri Jayawardenapura	Envirmt.Sc.
	Authors with Australia			
1	Senarathna, L.	18	University of Peradeniya	Medicine
2	Upul Senarath	18	University of Colombo	Medicine
3	Sheriff, M.H.R.	17	University of Colombo	Medicine
	Authors with Japan			
1	Tennakone, K.	19	Institute of Fundamental Studies	Physics
2	Wijayagunawardane, M.P.B.	18	University of Peradeniya	Agriculture
3	Kumara, G.R.A.	17	Institute of Fundamental Studies	Chemistry
	Authors with India			
1	Lakshman, P.L.N.	45	University of Ruhuna	Envirmt.Sc.
2	Sithambaresan, M.	26	Eastern University	Chemistry
3	Amerasinghe, P.H.	13	University of Peradeniya	Zoology
	Authors with Canada			
1	Premawardhena, A.	14	University of Kelaniya	Medicine
2	Karunaratne, V.	12	University of Peradeniya,	Chemistry
3	Samarasekera, R.	8	Industrial Technology Institute	Chemistry
	Authors with Sweden			
1	Dissanayake, M.A.K.L.	22	University of Peradeniya	Physics
2	Bandara, T.M.W.J.	14	University of Ruhuna	Physics
3	Jayasundara, W.J.M.J.S.R.	11	University of Peradeniya	Physics
	Authors with Germany			
1	Chandrajith, R.	17	University of Peradeniya	Agriculture
2	Weerasooriya, R.	14	Institute of Fundamental Studies	Agriculture
3	Dissanayake, C.B.	13	University of Peradeniya	Agriculture
	Authors with Thailand			
1	De Silva, H.J.	7	University of Peradeniya	Medicine
2	Amarasinghe, U.S.	7	University of Kelaniya	Zoology
3	Gunatilleke, N.	7	University of Peradeniya	Botany
	Authors with Netherland			
1	Ibrahim, M.N.M.	11	University of Peradeniya	Envirmt.Sc.
2	Rajindrajith, S.	8	University of Kelaniya	Medicine
3	De Silva, P.M.C.S.	5	University of Ruhuna	zoology

developing country like Sri Lanka in the field of research. King's College London (99 papers), University of Oxford (85 papers), University of New South Wales (74 papers) are the international institutions publishing the largest number of papers in collaboration with Sri Lanka.

The top 3 Sri Lankan scientists for collaborative publication with other countries and foreign scientists for collaborative publication with Sri Lanka are presented in tables 6 and 7. The top

Table 7.

Top 3 International prolific collaborators of Sri Lanka from each leading partners and their affiliating institute

No	UK Authors with SL	Output	Field	Affiliation
1	Sumathipala, A.	37	Medicine	Institute of Psychiatry
2	Siribaddana, S.	22	Medicine	Institute of Psychiatry
3	Hemingway, J.	19	Medicine	School of Tropical Medicine
	USA with SL			
1	Ashton, M.S.	15	Agriculture	Yale University
2	Thenkabail, P.S.	15	Agriculture	Yale University
3	Dias, H.V.R.	15	Chemistry	University of Texas
	Australia with SL			
1	Buckley, N.A.	60	Medicine	Canberra Clinical School
2	Eddleston, M.	53	Medicine	Canberra Clinical School
3	Dawson, A.H.	36	Medicine	Canberra Clinical School
	Japan with SL			
1	Fujimoto, Y.	22	Medicine	Toneyama National Hospital
2	Konno, A.	17	Medicine	Chiba University
3	Agatsuma, T.	14	Medicine	Univ.of Agri and Vet.Medicine
	India with SL			
1	Kurup, M.R.P.	23	Chemistry	Cochin University of S&Tech
2	Suresh, J.	46	Medicine	Reddy's Laboratory-Discovery Research
3	Vijayakumar, V.	14	Chemistry	VIT University
	Canada with SL			
1	Andersen, R.J.	12	Chemistry	University of British Columbia
2	Olivieri, N.F.	9	Medicine	University of Toronto
3	Carr, G.	8	Agriculture	National Water Research Institute
	Sweden with SL			
1	Cooray, V.	26	Physics	Uppsala University
2	Mellander, B.E.	22	Physics	Chalmers University of Technology
3	Albinsson, I.	13	Physics	Chalmers University of Technology
	Germany with SL			
1	Tobschall, H.J.	23	Chemistry	Johannes Gutenberg-Universität,
2	Eyer, P.	14	Medicine	Universität München
3	Worek, F.	8	Medicine	Universität München
	Thailand with SL			
1	Chongthaleong, A.	10	Medicine	King Chulalongkorn Memorial Hospital
2	Samarakoon, L.	9	Agriculture	Asian Institute of Technology
3	Thanuthong, T.	9	Agriculture	Songkhla Rajabhat University
	Netherland with SL			
1	Bastiaanssen, W.G.M	17	Agriculture	Land
2	Bos, M.G.	14	Agriculture	International Institute for Geo-information
3	Benninga, M.A.	8	Medicine	Emma Kinderziekenhuis ACM

Sri Lankan author is Lakshman,PLN as per the International Research Collaboration. He has recorded the highest number IRC. He was associated with the University of Ruhuna and contributed his maximum papers in collaboration with India. The top foreign author is Buckley, N.A. in terms of research publication with Sri Lanka. He was associated with Canberra Clinical School, Australia.

To get a better understanding of how collaborative papers are cited and in particular the high impact articles with foreign authors have the top 100 of the research papers of Sri Lanka according to their citations received were examined. The shocking fact that 97 of these papers had been published in collaboration with other countries was discovered. This shows to what extent Sri Lanka's research is dependent on International Research Collaboration in order to gain the impact of papers. Table 5 shows the top 5 highly cited articles with foreign collaboration. This table includes title of the article, author(s), citation count, name of the source and published year.

Table 8.

Top 5 highly cited collaborative papers

Citations	Title of Papers	Authors	Published	Source Title
409	Soil-transmitted helminth infections: Updating the global picture	De Silva N.R. et al.	2003	Trends in Parasitology
396	Dynamic response of dye-sensitized nanocrystalline solar cells: Characterization by intensity-modulated photocurrent spectroscopy	Ileperuma O. et al.	1997	Journal of Physical Chemistry B
369	Dynamic modeling of doubly fed induction generator wind turbines	Ekanayake J.B. et al.	2003	IEEE Transactions on Power Systems
362	Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: A systematic analysis for the Global Burden of Disease Study 2010	Dharmaratne S.D et al.	2012	The Lancet
362	User mobility modeling and characterization of mobility patterns	Zonoozi M.M., Dassanayake P.	1997	IEEE Journal on Selected Areas in Communications

Conclusion & Recommendations

International research Collaboration is often regarded as an effective way to get access to the advanced scientific inputs for the developing countries like Sri Lanka. In this view, research collaboration with the developed world is a means to reach the standard of research publication at international level. In addition, it is a yardstick to measure the exposure of researchers. The present changes that are being made in the higher educational system of Sri Lanka, pave the way for greater awakening in International Research Collaboration. The measures being taken by the UGC in the direction of promoting International Research Collaboration are a great source of encouragement to would-be researchers. We can safely expect a greater volume of International Research Collaboration in the future.

Financial constraints for R&D purposes restrict Sri Lankan scientists from competing with the rest of the world scientists. Sri Lanka cannot produce high quality research on its own due to increasing demand for collaborative research. To overcome these problems Sri Lanka has to get the collaboration, co-operation and support of other countries. Making use of these, collaborative and knowledge sharing with other countries can further enhance the chances of realizing our object to uplift the standard of research in Sri Lanka. In this regard, close co-operation and interaction with other nations would be the most appropriate course of action.

The study will provide research policymakers with a more complete picture of innovation capability in collaborative research, and help them to make better decisions. In addition, it will stimulate useful discussions among scientists and research managers, government and funding agencies about future research direction in Sri Lanka.

One of the important constituents for improvement of collaborative contribution and development in Sri Lanka is the initiative for academic, research and governance reforms in research institutions and higher education institutions as per the emerging age of research.

Collaborative activities primarily depend upon financial allocation. Therefore, funds raised from other national and global agencies, play an important role in aiding collaborative projects. In addition, incentives must be put in place to enable universities & R& D institutes of Sri Lanka to participate in international networks.

References

Adams,J (2013) The fourth age of research, *Nature,* 497 (30 May) 557-560.

Anuradha,K.T., Urs,S.R. (2007) Bibliometric indicators of Indian research collaboration patterns: a correspondence analysis, *Scientometrics* 71(2) 179–189.

Bookstein,A.,Moed,H., Yitzahki,M. (2006) Measures of international collaboration in scientific literature: part II, *Information Processing and Management*, 42, 1422–1427.

Freshwater,D.,Sherwood,G., Drury,V.(2006) International research collaboration: issues, benefits and challenges of the global network, Journal of Research in Nursing,11 (4) 295-303

Hafernik, J., Messerschmitt,D., Vandrick,S. (1997). Collaborative research : Why and How?. *Educational Researcher, 26 (9),31-35*

Kim,K. (2006). Measuring international research collaboration of peripheral countries: taking the context into consideration. *Scientometrics* 66 (2), 231-240.

Schmoch, U.,Schubert,T. (2008). Are international co-publications an indicator for quality of scientific research?, *Scientometrics,* 74 (3) 361 – 377.

Relationship between Economic Development and Intellectual Production

Umut Al and Zehra Taşkın

Hacettepe University, Department of Information Management, 06800, Beytepe Ankara, Turkey
umutal@hacettepe.edu.tr, ztaskin@hacettepe.edu.tr

Abstract

The level of economic development affects the design of different systems. At the country level, scientific outputs are related to the research and development expenditures. In this study, the relationship between economic development and intellectual production was investigated. The term "intellectual production" was used for the number of publications and patents. Patents were examined according to their types, which were national and triadic. Moreover, Research and Development (R&D) expenditures and Gross Domestic Products (GDP) were used as economic development indicators. In this study following research questions were addressed: 1. Is there any meaningful relationship between GDP and the number of patents? 2. Is there any meaningful relationship between GDP and the number of scientific publications? 3. Is there any correlation between R&D expenditures and patent production? 4. Is there any correlation between R&D expenditures and the number of scientific publications? In addition to these research questions, this paper focuses on the changes of economic development and intellectual production indicators throughout time. As a result, it was seen that countries show continuous improvement in years, both for economic development indicators and intellectual production indicators. Findings also showed that Luxembourg, USA, Switzerland, Norway and Israel are far beyond form other countries in terms of national income per person, Scandinavian countries distinctively separated from other countries especially in terms of the number of national patents per population and Switzerland, Sweden, Israel, Denmark and Finland share the first rows in the number of publications per population ranking.

Introduction

It is widely accepted that countries' scientific and technological progress and Research and Development (R&D) expenditures are related to the economic development levels. The measure of development level for countries, at first, had been natural capital, which is about the wealth of natural resources, rich oil deposits, fertile soils, etc. However, productivity of countries, which includes human capital, physical capital and natural capital, has been taking the place of pure natural capital (Soubbotina & Sheram, 2000, p. 11). The measurements of development levels are accepted as GDP (Gross Domestic Products) and GNP (Gross National Products). GDP is defined as the market value of goods and services produced within a selected geographic area (usually a country) in a selected interval of time (often a year). It is generally about outcomes rather than processes. Although GNP has similar meaning with GDP, multinational corporations are only calculated by GDP. GNP is a more local quantity (Leamer, 2009, p. 19).

Developed countries have large investments on R&D. At the same time, their scientific and intellectual production has been increasing year by year. The most important issues of science policy in each country are the structure and efficacy of R&D activity and its relation to GDP (Vinkler, 2007, p. 238). The challenges in cross-national comparisons of R&D expenditure and publication output were also reported in the literature (Wendt, Aksnes, Sivertsen & Karlsson, 2012). It is also mentioned that some factors, such as the coverage and comparability of countries in the *Web of Science*, differences in national research systems, may affect the validation and comparison (Wendt, Aksnes, Sivertsen & Karlsson, 2012, p. 830).

With this study, the relationship between the indicators of economic development (R&D expenditures, GDP) and intellectual production (number of national and triadic patents, and number of scientific publications) was investigated.

Literature Review

There are too many publications in the literature that point out the relationship between GDP and science and technology production and expenditures. One of the prior works about this relationship was written by Teitel in 1994 (Teitel, 1994). He used mathematical methods to calculate the relationship, and found statistically significant and meaningful results between patents, R&D expenditures, country sizes and per-capita incomes.

Ye's study (2007) found the strongest relationship between country development level and scientometric criteria. The correlation between GDP and scientific production was determined for 24 countries by using IMF, WIPO and UNESCO data. However, the author indicated that the results of study were based on only one year-data (2001) and further studies were needed to confirm these results.

Another study in the literature presented a model to test the relationship between R&D expenditures and number of patents, by evaluating case studies in the literature (Prodan, 2005). As a result, a strong positive correlation determined between R&D expenditures and patent applications was found. In addition to this, it was pointed out that the numbers of patents also differ from country to country.

A report (IDEA Consult, 2008) indicated that levels of R&D spending were interrelated to levels of economic growth. Findings showed that R&D intensities were temporarily influenced by the levels of GDP growth. However, the development patterns differ strongly among the countries depending on governance structure, policy priorities, and systematic features like industry and academic structures, which means "one size fits all" approach does not fit for all the countries.

A new indicator to analyse mean structural differences of different fields was found out in another study (Vinkler, 2007). A meaningful correlation was determined between GDP and number of publications in the longitudinal studies for countries. However, no direct relationship between GDP and information production of countries was found. It was noted in this study that R&D expenditures actually did not depend on real needs. However, one should note that, rich countries can always afford to spend more money on scientific research than poor countries.

Olwan (2013) focused on the correlation between intellectual property systems (IP systems) of countries and their development levels. This paper investigated developing countries from the point of effectiveness of their IP systems and its effects to their economies. As a result, it was found that there was no meaningful correlation between IP systems and economic development levels of developing countries.

Some studies in the literature concentrated on different effects of scientific outputs. In one of these studies (Nguyen & Pham, 2011), scientific output and its relationship with knowledge economy were examined in 10 South East Asian countries. This study (Nguyen & Pham, 2011, p. 113) found that there was a strong relationship between scientific output and knowledge economy index among the South East Asian countries. In a more recent study (Akhmat, Zaman, Shukui, Javed & Khan, 2014, p. 349), the empirical relationship between educational indicators and research productivity in top twenty nations of the world in terms of number of publications, citations and patents was examined. The results revealed that educational indicators were important to increase research productivity.

Many previous studies also found that there can be meaningful correlations between economic power and information production. However, it should not to be forgotten that these kinds of evaluations can change from country to country. The situation for The Organisation for Economic Co-operation and Development (OECD) countries is investigated by this study.

Methodology and Data Sources

This study analyses the related data belong to 34 OECD countries. The main aim of this research is to understand the relationship between the indicators of economic development and intellectual production. The term "intellectual production" in this study is defined as the number of publications and patents. Patents were also examined according to their types, which were national and triadic. "Economic development" indicators were identified as R&D expenditures and GDP. It would be interesting to see that whether different development levels of countries affect intellectual production. To achieve the aim of this paper, the following research questions are investigated:

- Is there any meaningful relationship between GDP and the number of patents (national and triadic)?
- Is there any meaningful relationship between GDP and the number of scientific publications?
- Is there any correlation between R&D expenditures and patent production?
- Is there any correlation between R&D expenditures and the number of scientific publications?

All of the OECD countries were selected to test correlations. GDP per capita, R&D expenditures, and number of patents data were gathered from *OECDiLibrary's National Accounts, Main Science and Technology Indicators* and *OECD Patent Statistics* databases (http://www.oecd-ilibrary.org). The number of scientific publications was collected from *Thomson Reuters' InCites*. The 34 members of OECD were very different than the others, in terms of population size. Therefore, all of the indicators were normalized according to population size. The population statistics were also obtained from OECD databases. The data of the study showed normal distribution after the normalization process, so the Pearson's correlation analysis was chosen for the correlation tests. To be able to use the Pearson's correlation analysis, median values of all indicators were calculated. Moreover, economic development and intellectual production indicators within a 30-years period (1981-2010) were also analysed within the scope of this study.

Findings

Today, OECD has 34 member countries (OECD, 2014). It was seen that within the 30-years period that we have dealt in this study, these countries progressed at various levels, in terms of the GDP, R&D expenditures, number of patents, number of scientific publications indicators. There is no doubt that this progress should be considered as normal. However, the number of publications had been increased enormously within the years (Figure 1 & Table 1). Although this can be based on the regional development policy of citation indexes (Testa, 2008), there can be also different reasons for each country.

In Table 1, the number of publications of the 34 countries for a six period of five-year intervals was given. These numbers showed that the number of publications of the countries has been increasing over time. Although such amount of increase was not observed, it is known that the number of patents was also escalating. In addition to this, a substantial increase in the share of R&D expenditures and national incomes of the countries was also recognized. To be able to make meaningful comparisons, population information of the countries was used.

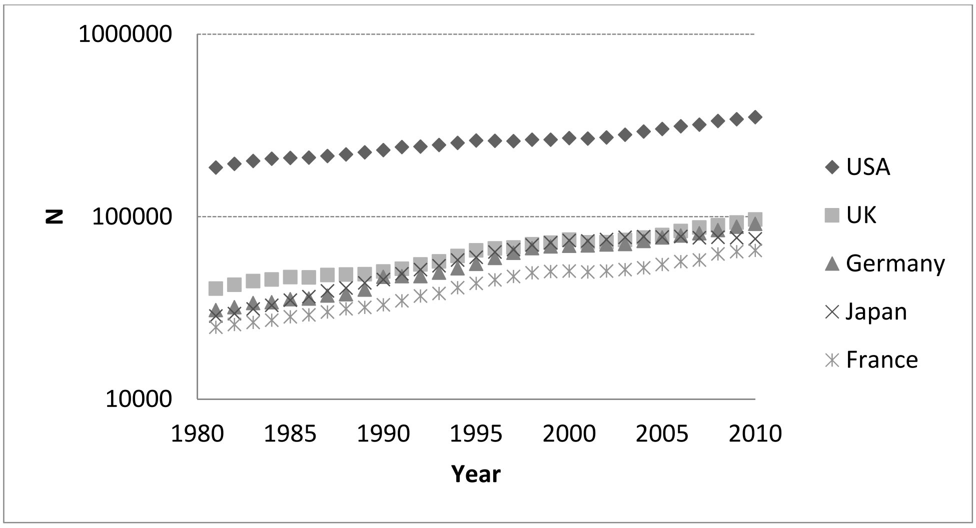


Figure 1. The five most productive countries by year

Table 1. Number of publications by periods

	Periods					
Countries	*1981-1985*	*1986-1990*	*1991-1995*	*1996-2000*	*2001-2005*	*2006-2010*
Australia	57,003	64,710	80,299	105,033	124,051	178,197
Austria	15,721	18,099	23,432	33,888	42,152	52,999
Belgium	24,162	27,977	36,265	48,874	59,658	78,829
Canada	116,329	142,088	165,738	173,180	194,304	261,703
Chile	3,730	4,970	6,366	8,580	12,968	21,299
Czech Republic	3	221	7,678	20,804	26,852	40,551
Denmark	20,498	23,258	29,680	37,546	42,780	53,720
Estonia	4	19	1,112	2,683	3,376	5,621
Finland	15,909	19,060	25,802	34,855	40,435	47,907
France	132,255	155,129	193,356	241,844	258,656	307,133
Germany	165,666	197,023	251,162	328,050	361,529	423,944
Greece	6,062	9,240	14,347	22,070	33,159	50,123
Hungary	14,676	14,746	15,200	19,193	22,824	27,225
Iceland	362	557	1,038	1,457	2,077	3,243
Ireland	5,389	6,451	8,539	12,662	17,187	28,684
Israel	29,505	33,540	39,046	47,096	52,838	59,191
Italy	59,818	78,038	112,544	151,205	186,869	243,143
Japan	156,819	205,040	271,717	346,284	381,107	383,844
Korea	2,025	5,766	17,592	52,950	105,304	171,983
Luxembourg	132	131	253	408	713	1,866
Mexico	5,344	7,108	11,590	21,327	31,172	43,782
Netherlands	43,218	57,910	77,438	94,728	108,303	141,569
New Zealand	12,137	13,334	15,905	21,379	24,551	33,165
Norway	13,247	14,676	19,136	24,041	28,787	42,801

Poland	24,621	28,511	32,039	44,412	63,356	89,239
Portugal	1,700	3,341	6,554	12,948	22,790	39,335
Slovak Republic	2	71	4,535	10,366	10,377	13,894
Slovenia	1	25	2,326	6,064	9,129	14,845
Spain	24,180	41,269	67,972	104,109	136,859	202,237
Sweden	40,932	49,107	59443	74,151	81,914	95,030
Switzerland	36,058	40,301	52635	67,649	77,618	102,996
Turkey	2,014	3,865	9175	22,249	53,971	97,619
UK	219,062	241,188	289777	352,238	375,505	450,002
USA	1,000,825	1,102,604	1245611	1,318,469	1,416,532	1,660,017

In Figure 2, the relationship between the number of publications per 1000 people and GDP per head was shown in the country level. The numbers in the Figure represents the median values of the 30-years data. Scandinavian countries (such as, Sweden, Denmark, Finland) were recognized in the Figure, in terms of both for the number of publications per 1000 people and GDP per head. Along with these countries, Switzerland and Israel were also came to the forth, in terms of the number of publications per population. On the other hand Luxembourg, which has the highest national income, located in the bottom of the list in terms of the number of publications per population, like Mexico, Turkey, Chile and Korea. In general, it was observed that, the countries which have the highest number of publications per 1000 people have also the highest GDP per head.

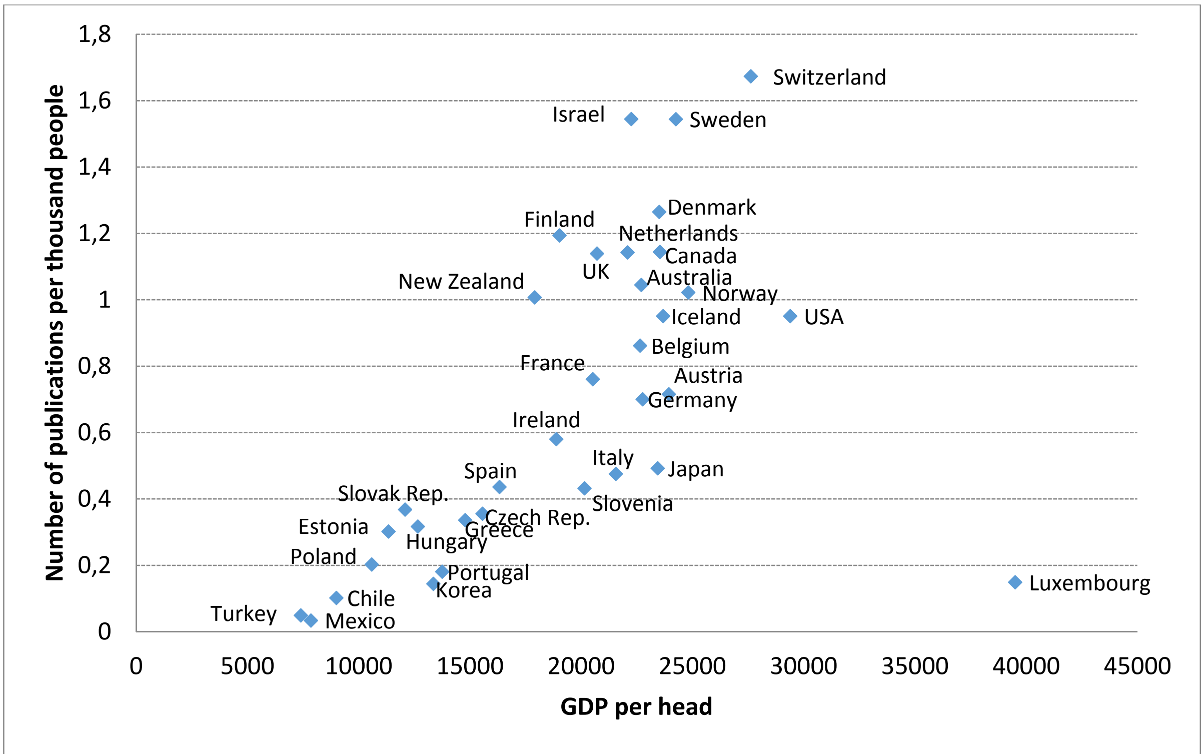


Figure 2. Number of publications per thousand people and GDP per head

Switzerland, Japan, Sweden, Germany, Netherlands, USA and Finland are leading countries in terms of the number of triadic patent per country population. Turkey and Mexico has the worst performance in terms of the number of triadic patents per million population along with Estonia, Chile, Poland, Slovak Republic, Portugal, Greece and Czech Republic. These countries have less than one triadic patent per million population (Figure 3).

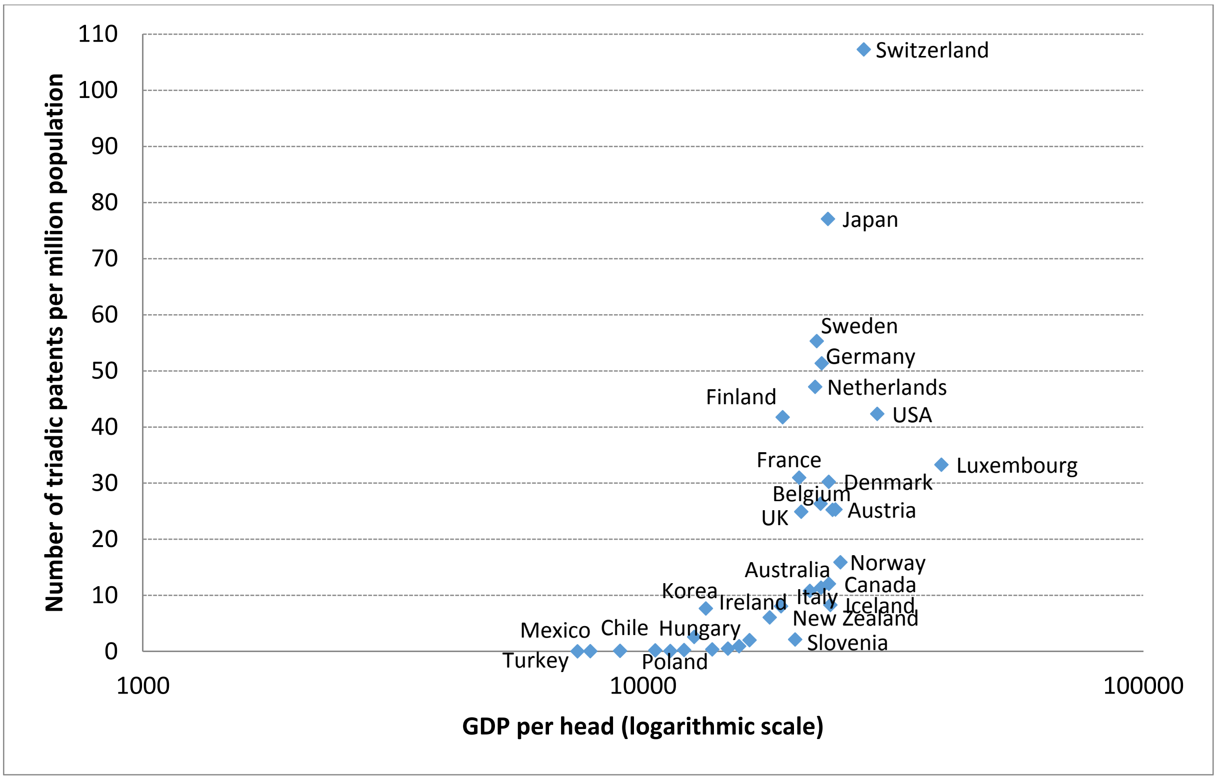


Figure 3. Number of triadic patents per million population and GDP per head

Luxembourg, Israel, Sweden, USA and Japan are the first five countries that have the highest R&D expenditures per person (Figure 4). The general trend shows that the countries (such as, Mexico, Turkey, Chile) that has limited shares for R&D expenditures has also the lowest numbers of publications per population.

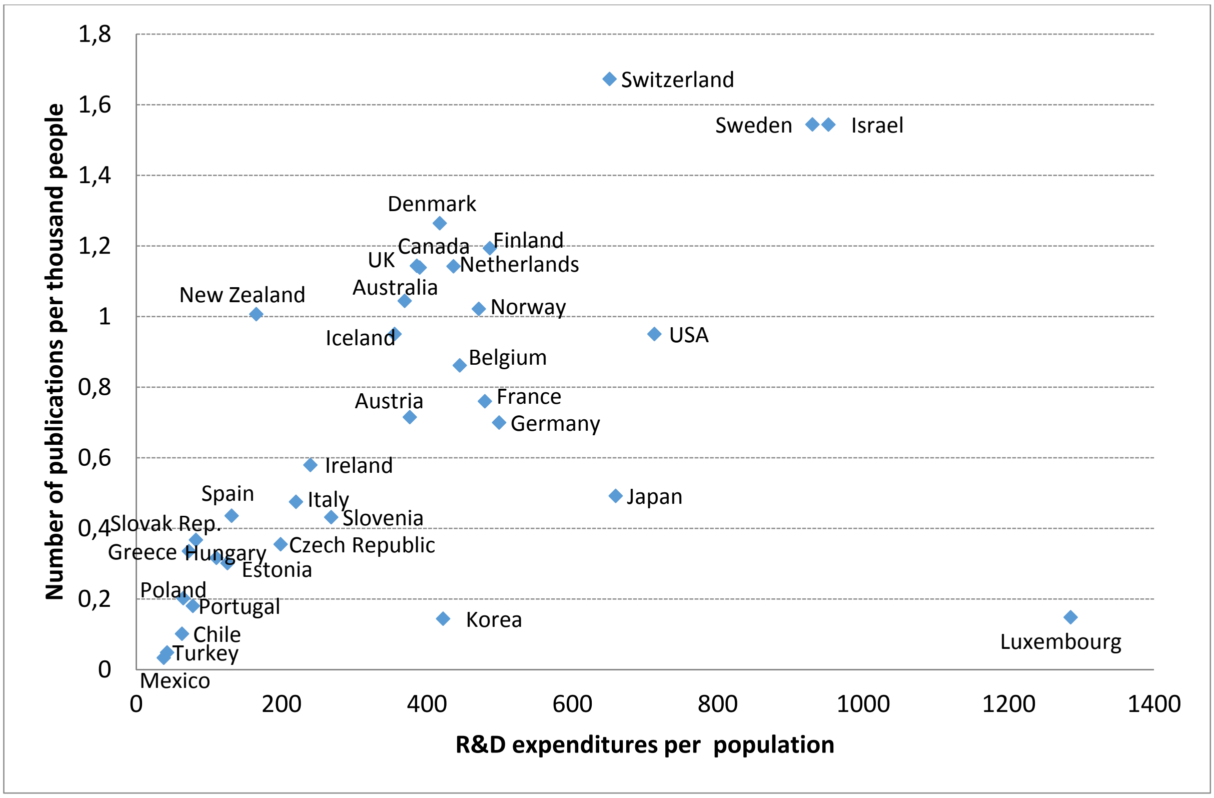


Figure 4. Number of publications per thousand people and R&D expenditures per population

In this study the values that were obtained from the division of R&D expenditures to triadic patent numbers were also compared. In other words, we tried to see the amount of money that the countries spend for triadic patents. It was found out that, Switzerland, Japan, Netherlands, Germany and Finland are the most remarkable countries in terms of the ability to transform the R&D expenditures to patents. Contrary to this, it was identified that Estonia, Turkey, Chile, Mexico and Poland are the ones which spent the most money to have a triadic patent. Figure 5 revealed that R&D expenditures per population and the number of triadic patents per million population are similar to each other.

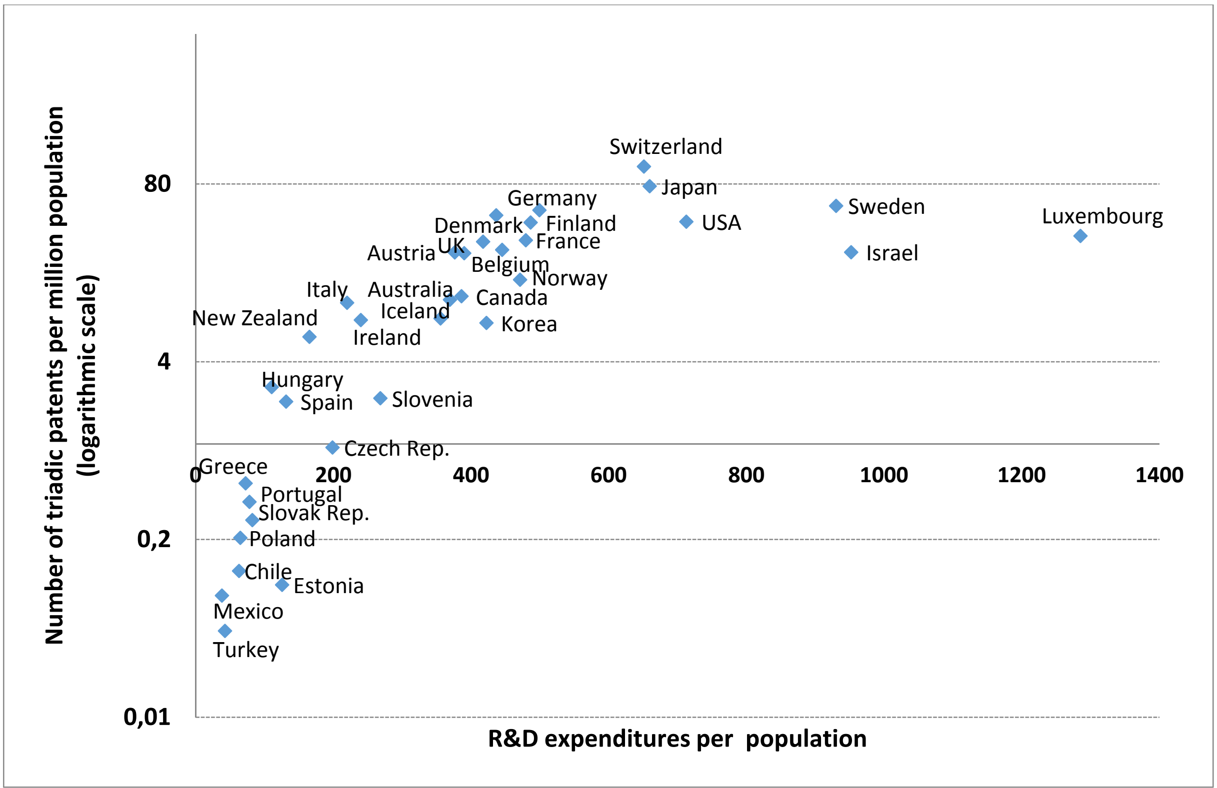


Figure 5. Number of triadic patents per million population and R&D expenditures per population

Some statistical tests were conducted on the raw data that forms Figure 2, 3, 4 and 5. According to this, significant correlations were observed among all of the economic development indicators and all of the intellectual production indicators (Table 2).

Table 2. Pearson correlation coefficients among variables

	Intellectual production indicators		
Economic development indicators	*Number of publications per population*	*Number of triadic patents per million population*	*Number of national patents per million population*
GDP per head	0.561	0.604	0.567
R&D expenditures per population	0.524	0.667	0.674

Note: Correlations are significant at the 0.01 level.

In the light of the statistical evaluations, the answers of our research questions are as follows:

- There is a positive correlation between GDP per head and the number of publications per population was statistically significant (Pearsons's $r = .561$, $p < .01$).

- There is a positive correlation between GDP per head and the number of triadic patents per million population was statistically significant (Pearsons's $r = .604$, $p <.01$).
- There is a positive correlation between GDP per head and the number of national patents per million population was statistically significant (Pearsons's $r = .567$, $p <.01$).
- There is a positive correlation between R&D expenditures per population and the number of publications per population was statistically significant (Pearsons's $r = .524$, $p <.01$).
- There is a positive correlation between R&D expenditures per population and the number of triadic patents per million population was statistically significant (Pearsons's $r = .667$, $p <.01$).
- There is a positive correlation between R&D expenditures per population and the number of national patents per million population was statistically significant (Pearsons's $r = .674$, $p <.01$).

Conclusion

The relevant investments show the importance given by the countries to science and R&D. There is no doubt that scientific productivity level of the countries is affected by not only qualified manpower but also by the economic development levels. Today, parallel to the increase of the importance of knowledge as an economic value, it is witnessed that most of the countries increase their investments for the production of theoretical knowledge which is aimed to be transformed to product and services and develop some new policies towards this goal. Patents and scientific publications, which are the products of labour-intensive work, clearly show the level of investments of the countries for science and scientists. From this point, it can be said that competitive advantage of the countries is also parallel to their productivity level of information. From a systems approach perspective, when we take the expenditures as an input, the cost of patents and scientific publications, which can be counted as outputs, must be questioned.

With this research, it is seen that most of the "rich countries" make some important contributions to the world literature, in terms of publications and patents. In this context, some results of this study are similar to the literature. The correlation was observed between GDP per head, R&D expenditures per population, number of publications per population, number of national patents per million population and number of triadic patents per million population.

Country-based findings that we obtained from this research are as follows:

- It was seen that countries show continuous improvement in years, both for economic development indicators and intellectual production indicators.
- Luxembourg, USA, Switzerland, Norway and Israel are far beyond the OECD countries such as Hungary, Slovak Republic, Estonia, Poland, Chile, Mexico and Turkey in terms of national income per person. Similar situation is observed for the R&D expenditures of the countries. R&D expenditures of Luxembourg, Israel, Sweden, USA and Japan per person are 10 to 25-fold higher than that of Greece, Poland, Chile, Turkey and Mexico.
- Scandinavian countries such as Sweden, Finland, Denmark, distinctively separated from other countries especially in terms of the number of national patents per population. Switzerland and Japan are two leading countries in terms of the number of triadic patents per population.
- Switzerland, Sweden, Israel, Denmark and Finland share the first rows in the number of publications per population ranking, where Luxembourg, Korea, Chile, Turkey and Mexico are in the bottom among the 34 OECD countries.

Such topics like, the effect of the number of researchers in the countries to the number of research outputs, the contribution of the universities to the national intellectual production, the effects of patents to science, technology and innovation policies of the countries can be considered for the future studies.

References

Akhmat, G., Zaman, K., Shukui, T., Javed, Y. & Khan, M.M. (2014). Relationship between educational indicators and research outcomes in a panel of top twenty nations: Windows of opportunity. *Journal of Informetrics*, 8(2), 349-361.

IDEA Consult. (2008). A time series analysis of the development in national R&D intensities and national public expenditures on R&D. *Final Study Report for Specific Assignment 4 (SPA4): Literature Review, Feasibility Study, Collection of Field Statistics, Calculation of Indicators, Survey, Conclusions*. Brussels. Retrieved June 1, 2014 from: http://ec.europa.eu/invest-in-research/pdf/download_en/spa4_final_report_final.pdf

Leamer, E.E. (2009). *Macroeconomic patterns and stories: A guide for MBAs*. Heidelberg: Springer.

Nguyen, T.V. & Pham, L.T. (2011). Scientific output and its relationship to knowledge economy: an analysis of ASEAN countries. *Scientometrics*, 89(1), 107-117.

OECD. (2014). Members and partners. Retrieved June 1, 2014 from: http://www.oecd.org/about/membersandpartners/

Olwan, R.M. (2013). Intellectual property and economic development. In *Intellectual Property and Development* (pp. 99-151). Berlin: Springer Verlag.

Prodan, I. (2005). Influence of research and development expenditures on number of patent applications: Selected case studies in OECD countries and Central Europe, 1981-2001. *Applied Econometrics and International Development*, 4-5, 5-22.

Soubbotina, T.B. & Sheram, K.A. (2000). *Beyond economic growth: Meeting the challenges of global development*. Washington DC: World Bank.

Teitel, S. (1994). Patents, R&D expenditures, country size, and per-capita income: An international comparison. *Scientometrics*, 29(1), 137-159.

Testa, J. (2008). *Regional content expansion update: Web of Science 5.0.* Retrieved June 1, 2014 from: http://wokinfo.com/media/pdf/RExEssay.pdf

Vinkler, P. (2007). Correlation between the structure of scientific research, scientometric indicators and GDP in EU and non-EU countries. *Scientometrics*, 74(2), 237-254.

Wendt, K., Aksnes, D.W., Sivertsen, G., & Karlsson, S. (2012). Challenges in cross-national comparisons of R&D expenditure and publication output. *Proceedings of 17th International Conference on Science and Technology Indicators* (pp. 826-834). Montréal: Science-Metrix and OST.

Ye, F.Y. (2007). A quantitative relationship between per capita GDP and scientometric criteria. *Scientometrics*, 71(3), 407-413.

The Impact of Turkey in the Library and Information Science Literature

Umut Al*, İrem Soydal*, Umut Sezen** and Orçun Madran*

*Hacettepe University, Department of Information Management, 06800 Beytepe, Ankara, Turkey
umutal@hacettepe.edu.tr, soydal@hacettepe.edu.tr, omadran@hacettepe.edu.tr

**Hacettepe University, Department of Electrical and Electronics Engineering, 06800 Beytepe, Ankara, Turkey
u.sezen@ee.hacettepe.edu.tr

Abstract

This paper investigates the contribution of Turkey to the world library and information science literature. In this study we investigate the bibliometric characteristics of 219 library and information science journal articles written by authors affiliated with Turkish institutions and indexed in the *Social Sciences Citation Index* (*SSCI*) between the years 1974-2013. The most preferred library and information science journals to publish articles for Turkey addressed authors are Information Processing & Management and Scientometrics. 159 of 219 articles of Turkey addressed library and information science articles were cited at least once by the publications in citation indexes. Total number of citations of Turkey addressed 159 articles was 1304. All of the contributions were written in English and one-third of them had single authorship. Turkey addressed library and information science articles have been cited by 69 different countries and in particular mostly by the United States of America, China and England. On the other hand the country self-citation was 27%.

Introduction

The importance of publishing articles in journals within the scope of citation indexes has been increasing gradually. Researchers wish to publish their articles in the journals that are indexed in citation indexes to be able to spread their studies to wide audiences. At the same time, it is a known fact that publishing articles within the scope of citation indexes usually counted as an indicator of reputation. On the other hand, it is also a motivating effect for academicians in terms of tenure and promotion. Regardless of the reason, it is a fact that, both the number of researchers, who want to publish their works in the journals that are indexed in the citation indexes, and articles published in these indexed journals increased in the recent years. The same situation is also observed in the library and information science (LIS) literature in particular.

Turkish interest in library science, (especially in terms of LIS education) can be traced back to the invitation of John Dewey who accepted to carry out a survey on Turkish education system and make recommendations. His 1924 dated report included a suggestion for some Turkish experts to be sent to the United States for librarianship training. Some courses on library practice were offered in İstanbul and Ankara until 1953 (Whitten & Minder, 1974, pp. 223-224). In 1954 Ankara University Institute of Librarianship was established as the first LIS Bachelor's program in Turkey followed by İstanbul University and Hacettepe University (Kum & Erdoğan, 1980). Today, six information management departments actively educate the future LIS professionals. Moreover, these departments also contribute the academic literature of LIS both in Turkey and in the world. With this study, we tried to identify the contribution of Turkey to the world LIS literature.

Literature Review

It is seen in the literature that researchers examine not only the bibliometric characteristics of a particular LIS journal (Bonnevie, 2003; Furner, 2009; Ginn, 2003; Mukherjee, 2009; Ramesh & Nagaraju, 2000; Schubert, 2002; Tsay & Shu, 2011), but also some of them compare more than one journal in terms of their bibliometric features (Harter, Nisonger & Weng, 1993; He &

Spink, 2002; Kajberg, 1996; Kim, 1991). It would be better to note that some of the aforementioned studies especially focused only on the citation analysis.

In one of the studies, the co-citation rate of *Journal of Information Science* was investigated (Bonnevie, 2003). The similarity measures based on co-citation analysis showed that the *Journal of Information Science* and *JASIS* were the closest ones in terms of their position in the LIS network (Bonnevie, 2003, p. 20). In another study (Schubert, 2002), a statistical overview of the first 50 volumes of the journal *Scientometrics* in its first 24 years was given. The study revealed the references that were highly cited by the articles published in *Scientometrics*, as well as the most citing references of *Scientometrics* that were cited by the articles published in the other journals. According to this, the most cited publication in *Scientometrics* articles was De Solla Price's *Little Science Big Science* (Schubert, 2002, p. 14).

Studies related to LIS journals covered both long and short term periods for the analysis. For instance, Ginn (2003) analysed relatively a small number of citation data that belongs to a short term period. The study revealed that the most cited journal was *JASIS*. One of the hypotheses of the study, which was defined as "the use of Web sites as reference sources would increase in number from the first year to the next", was not supported by the findings (Ginn, 2003, p. 108). Another study, (Harter, Nisonger & Weng, 1993) related to librarianship journals, examined the articles in three different journals (*College and Research Libraries*, *JASIS*, *Library Journal*) and investigated the semantic relationship between citing and cited documents in these journals to reveal the differences between them. For example, a clear difference was observed among the three journals regarding the type of the references they cite. A chi-square test showed that *JASIS* cites far more proceedings than expected. By contrast, *College and Research Libraries* and *Library Journal* cite books more than *JASIS* (Harter, Nisonger & Weng, 1993, pp. 545-546).

Another published study (Kajberg, 1996) reported the geographic distribution of cited documents in Danish that were published in joint Nordic LIS journals. The findings showed that about two thirds of cited documents were published in Denmark (Kajberg, 1996, p. 77). Similar findings were revealed by other researchers as well. In one of these studies, the journal named *Indian Journal of Information, Library and Society* was examined and it was observed that 60% of cited journals were from India (Ramesh & Nagaraju, 2000, p. 177). These studies point out the high usage rates of domestic publications in the LIS literature.

There are some studies in the literature (Herrero-Solana & Ríos-Gómez, 2006; Schloegl & Stock, 2004) that are similar to our study in which LIS journals that are included in the *Journal Citation Reports* (*JCR*) classification were investigated. One of these studies (Herrero-Solana & Ríos-Gómez, 2006) examined the contribution of Latin America between the years 1966 and 2003 by using the information science journals in the *JCR*. The study revealed that Brazil was the most productive country since it produced half of the total number of publications. Another study (Schloegl & Stock, 2004) investigated nearly 90.000 citations in 50 journals included in the *JCR*. The study compared the data obtained from citations and reader survey analysis. The relationships between different variables, such as reading frequency and impact factor; reading frequency and half-life and reading frequency and the number of references per article were tested. It was found out that the reading behavior was not affected by the impact factor (Schloegl & Stock, 2004, p. 1166).

Another bibliometric study on information science journals (Uzun, 2002) examined the articles in 21 journals (which were defined as core information science journals by the author) without a classification based on *JCR*. The impact of the publications of East European and developing countries on information science were discussed in the study according to a bibliometric approach. The study revealed that the above mentioned countries had limited publications on information science between the years 1980 and 1999 and only 8% of all articles were written by the authors from these countries (Uzun, 2002, p. 21).

In Turkey, bibliometric studies on LIS journals were based on a journal, titled the *Turkish Librarianship*, which has almost a 60-year history. In one of the earliest studies on bibliometrics in Turkey (Çakın, 1980), content analysis was done for the *Bulletin of the Turkish Librarians' Association* (the old title of the *Turkish Librarianship*). The first citation analysis study was conducted in 1996 and investigated the citations of articles of the *Turkish Librarianship* journal published during 1981-1995. The study has examined only the journal citations and found out that the most cited journal was the *Turkish Librarianship* (Kurbanoğlu, 1996a, p. 109). In another study, the accuracy of the citations appeared in the articles of the *Turkish Librarianship* during 1991-1995 was analysed (Kurbanoğlu, 1996b).

Another study aimed to determine the number of citations of refereed and non-refereed articles and the number of articles that were written in Turkish and in other languages, published in the *Turkish Librarianship* journal between the years 1992-2001 (Gürdal, 2002). The most comprehensive study on the *Turkish Librarianship* journal covers the years 1987-2001 (Tonta, 2002). The study revealed the results of extensive citation analysis of the articles published in the Turkish Librarianship journal and introduced some bibliometric information about articles that were published in the journal.

Different from the above mentioned ones, another study evaluated the citations of articles that were published in LIS journals by the researchers of Library and Information Science Departments of three different universities. The results of this study showed that these researchers tended to cite the articles of their fellow academicians who work in the same department (Yılmaz, 2000).

Methodology and Research Questions

Our study covers the years 1974 (which is the publication year of first Turkey addressed information science article in *SSCI*) and 2013 and it is intended to answer the following research questions:

- In which journals do Turkish LIS scholars publish more often?
- How many articles, which are indexed in the citation indexes, are produced by the scholars of LIS discipline?
- What percentage of articles authored by Turkish LIS scholars receive citations?
- What are the country origins of the authors who cited the Turkey addressed information science articles?

Data of this study come from Thomson Reuters' *Web of Science* online database. We searched *SSCI* on February 20, 2014 to identify the Turkey addressed articles published in LIS journals. To obtain the data, journal names were entered in "publication name" field and Turkey was entered in the "address" field. Journal names were obtained from Information Science & Library Science subcategory within the *JCR* 2012. *JCR* 2012 contains 85 journals in the related field (Thomson Reuters, 2013). The former *JCR*s were also examined because of the possible existence of some other journals related to this field in the previous years.

It was seen that a few of the journals have changed their names over time. To be able to make accurate evaluations, changes in the names of journals were determined and all the data belong to the ones that changed their names were classified under their new names.

For better interpretation of our study, there are three points which should not be overlooked. Firstly, in our study instead of considering if the authors' nationality was Turkish or not, it was examined whether the authors produced Turkey addressed articles. The second important point is that our research was based on the classification of Thomson Reuters. Although there are many journal titles in the Thomson Reuters' classification, it should be taken into account that some journals within this classification are directly related to the information science where the others have indirect associations with the field. Thirdly, it should be kept in mind that researchers from information science discipline in Turkey have also some publications in the journals that are not counted as information science journals. Naturally, these articles were not addressed within the scope of this research.

Findings and Discussion

The Turkey addressed information science articles indexed within the scope of *SSCI* between the years 1974-2013 are 219. In the same time period the number of information science articles from all over the world is 89,407. Hence, it is observed that the contribution of Turkey to the world information science literature is 0.24%.

Figure 1 shows the contribution of Turkey to the information science literature in five years periods. These periods display the proportion of the number of Turkey addressed articles to the number of whole articles in information science literature published within the scope of *SSCI*. In the first four periods (1974-1978; 1979-1983; 1984-1988 and 1989-1993) the contribution of Turkey to the world's information science literature was ten in the ten-thousand. The ratio accelerated in the following years and a dramatic increase was observed in the 2004-2008 period, which was fifty-seven in the ten-thousand (0.57%). The increase in the number of Turkey addressed information science articles seems to be higher than that of the other time periods. If this increase continues, it is estimated that Turkey addressed information science articles will take place in the world information science literature with more than one percent.

The deceleration which appears to be in the 2009-2013 period (see Figure 1) is related to the indexing process of citation indexes. As of February 2014, all articles that were belong to the year 2013 have not been indexed yet.

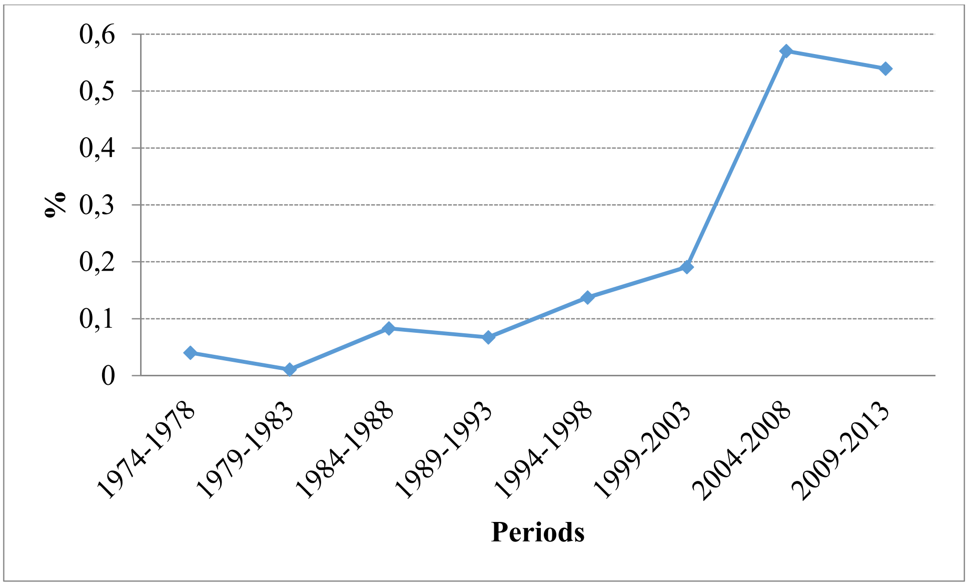


Figure 1. The percentage of Turkey addressed information science articles in the *SSCI*

In this research, nearly 100 information science journals were analyzed in order to investigate whether they have Turkey addressed articles or not. It was found out that Turkey addressed articles were published in 49 different information science journals. In Table 1, the names of the information science journals in which Turkey addressed articles took place were presented. Accordingly, more Turkey addressed articles were published in *Information Processing &*

Management and *Scientometrics*, followed by *Journal of the American Society for Information Science & Technology*, *Libri* and *Journal of Academic Librarianship*. Forty-one percent of the total number of Turkey addressed articles in the information science literature came from these five journals.

Table 1. Journals publishing eight or more contributions by Turkish institutions

Journal	*N*
Information Processing & Management	25
Scientometrics	25
Journal of the American Society for Information Science & Technology	18
Libri	13
Journal of Academic Librarianship	10
Government Information Quarterly	9
Journal of Information Science	9
International Information & Library Review	8
Telecommunications Policy	8
Other journals	94
Total	219

Until the date we conducted the search for this study, 219 Turkey addressed articles that were published in the information science journals received 1304 citations in total and 60 of them did not receive citations. However, 20 of these 60 articles published in 2013, therefore some more time is needed in order for them to get citations. The oldest article that did not receive citations was published in 1978. Six articles that were published in the years 1990-1998 and 33 articles that were published in the years 2002-2012 received no citations at all. The most frequently cited publication was cited 77 times. The average number of citations received by Turkey addressed articles was six.

The total number of different authors contributing to 219 articles was 341. Eighty percent (273 authors) of all authors contributed to the literature with only a single publication. There were nine authors publishing seven or more contributions in LIS journals.

It is also investigated that how many of the articles were produced by the authors who were working in information science discipline. 64 out of 219 articles (29%) were produced by the aforementioned scholars or graduates of information science discipline. In other words, it is seen that researchers out of information science discipline are more productive in the field. The most productive two researchers Seda Özmutlu and Hüseyin Cenk Özmutlu (13 and 12 articles, respectively) were industrial engineers. One of the scholars from Statistics department, Ali Uzun, is also one of the most productive authors in the information science discipline (9 articles). The most productive researcher from information science discipline is Yaşar Tonta (with 13 articles), who still is a faculty member in the Department of Information Management of Hacettepe university.

Publications with multiple authors constituted two-third of all articles. 33% of articles has two authors, 19% of articles has three authors, and 14% of articles has four or more authors.

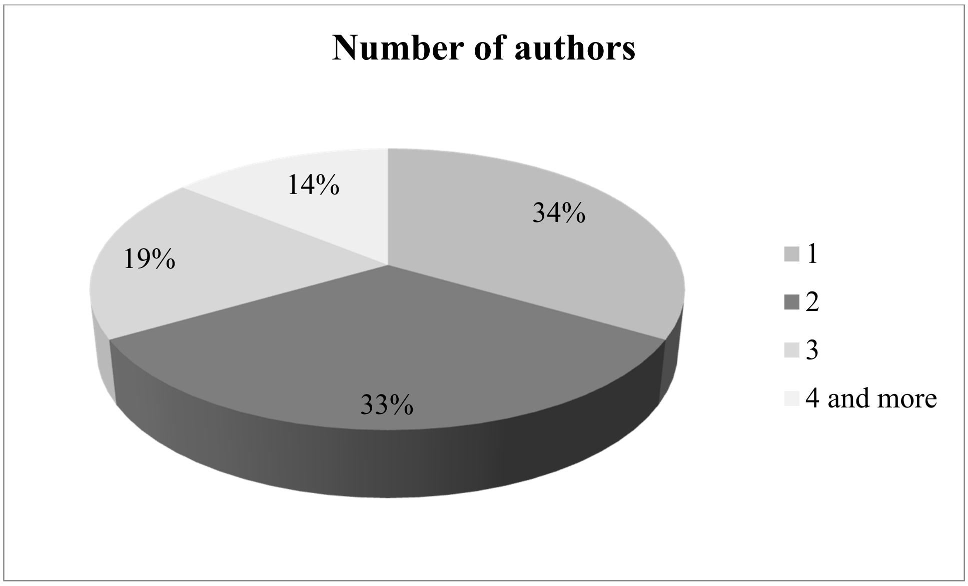


Figure 2. Number of authors

The average number of authors per article was two. Nevertheless, we observed a tendency towards multiple authorship (Figure 3).

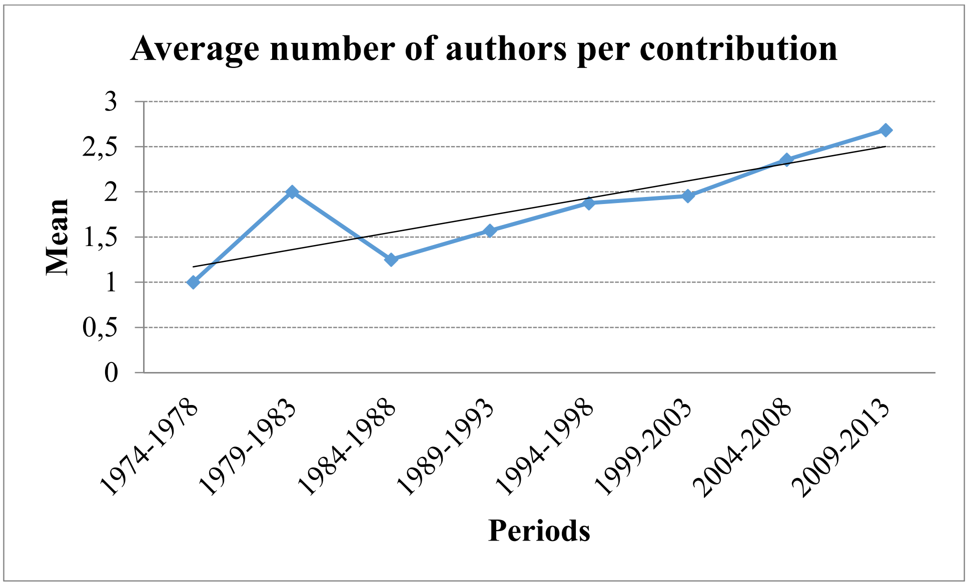


Figure 3. Average number of authors per contribution

Turkey addressed information science articles has received citations from 548 different publications, in which proceedings books were also included. More than one third of these citations come from 18 different journals. These journals and the number of citations that were received by Turkey addressed articles were shown in Table 2. Turkey addressed articles had mostly received citations from the articles that were published in *Journal of the American Society for Information Science & Technology*, *Scientometrics* and *Information Processing & Management*. These three journals, which were also addressed in Table 1, were information science journals where Turkey addressed articles published the most. In addition to these findings, the rank order correlation between the lists of publishing journals and cited journals was statistically significant (Spearman's *rho* =.621, p <.01).

Table 2. The most heavily citing journals for Turkey addressed information science articles

Journal	*N*
Journal of the American Society for Information Science & Technology	91
Scientometrics	84
Information Processing & Management	58
Online Information Review	32
Experts Systems with Applications	31
Journal of Academic Librarianship	20
Journal of Documentation	19
Journal of Information Science	19
Telecommunications Policy	19
Information Research	17
Government Information Quarterly	15
Library & Information Science Research	15
Libri	13
Electronic Library	12
Information & Management	12
Aslib Proceedings	11
International Information & Library Review	10
Program	10
Total	488

Turkey addressed information science articles had also received citations from the journals (such as, *Expert Systems with Applications*, *Simulation Modelling Practice & Theory*), which are classified under the disciplines that are not related to information science field (such as, education, mathematics) in terms of *JCR* classification. This shows, in a sense, the interest of different disciplines on information science field and reinforces the interdisciplinary nature of the field.

Table 3. The most citing countries for Turkey addressed information science articles

Country	*N*
Turkey	354
USA	241
China	108
England	95
Spain	72
Taiwan	64
Australia	53
Canada	48
India	37
South Korea	29
Netherlands	25
Malaysia	24
France	23
Germany	23
Iran	22

In this study, the countries that cited Turkey addressed information science articles were also investigated and findings showed that these articles had received citations again from the

Turkey addressed scholars. 354 out of 1304 citations (27%) were country self-citations, followed by United States of America, China and England. The countries which cited Turkey addressed articles were presented in Table 3 and in Figure 4 show the density of consideration of the countries worldwide.

Figure 4 shows that, although they are small in number, Turkey addressed information science articles were visible for and cited by several different countries. Argentina, Bangladesh, Botswana, Bulgaria, Costa Rica, Iceland, Kuwait, Lebanon, Lithuania, Mozambique, Peru, Philippines, Qatar, Romania and Tanzania are some of the countries which cited Turkey addressed information science articles. There are a small number of countries which have never been cited these articles and this reveals that some of these articles visible for a wide geographical area.

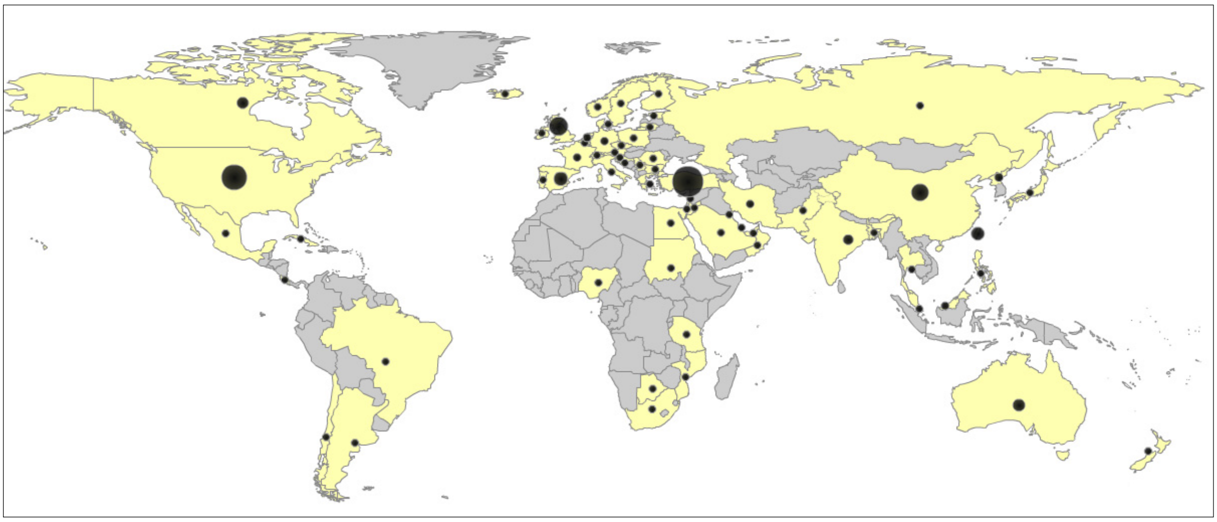


Figure 4. Citation map of Turkey addressed information science articles

Conclusion

Most of the information science articles in the citation indexes are west-originated, in particular United States of America and England. Scholars in Turkey also work for finding themselves a place in the world's information science zone. It can be said that the contribution of Turkey, in terms of publication, to information science field has increased over time. Different types of encouragements in Turkey's academic environment also had an impact on the motivation of scholars to produce articles which will be indexed in the citation indexes. The conclusions that have arisen from this study, which we tried to reveal the contribution of Turkey to information science literature of the world, were listed below:

- When the total number of journals in the *JCR* is taken into account, it is seen that Turkey did not contribute to most of these journals, conversely a few journals were focused on to publish articles. This situation can negatively affect the visibility of the Turkey addressed articles.
- It is observed that people who have not been working in the information science field show quite a serious attention to information science journals. The interdisciplinary nature of the field, as well as the academic promotion criteria in Turkey causes this. It is thought that as the number of information science departments in Turkish universities increase, the number of Turkey addressed articles will increase as well.
- One fourth of Turkey addressed information science articles has never been received citations. The most important reason for this is not enough time has passed for these articles in order for them to receive citations. Although it does not show a normal distribution, it is found out that Turkey addressed information science articles have been

cited an average of six. This is important, because the increase in Turkey's contribution to information science field is also observed in the increase of the citations to these publications.

- Turkey addressed information science articles used by the researchers from 69 different countries. Density of citations showed that the countries produce more articles (USA, England and China) cited more Turkey originated articles. Moreover, the spread of Turkey addressed articles to different countries in the world reveals that the limited contribution of this country, at least, had an effect on a wide geographical area.

In this study the Turkey's contribution to international information science literature, in terms of articles, were examined from various aspects. Looking from the international perspective, it is a known fact that not only having publications indexed in the citation indexes but also to have the journals covered by those indexes is important. Although Turkey has some essential journals on information science, they have not been indexed by the citation indexes yet. To make the articles published in the journals in Turkey more widely accessible, and to increase the visibility of these publications within the international information science literature, some attempts must be done towards making these journals covered by the citation indexes.

References

Bonnevie, E. (2003). A multifaceted portrait of a library and information science journal: the case of the Journal of Information Science. *Journal of Information Science*, 29(1), 11-23.

Çakın, İ. (1984). Turkish libraries: Historical context. *International Library Review*, 16(1), 71-77.

Furner, J. (2009). Forty years of the Journal of Librarianship and Information Science: A quantitative analysis, Part I. *Journal of Librarianship and Information Science*, 41(3), 149-172.

Harter, S.P., Nisonger, T.E. & Weng, A.W. (1993). Semantic relationships between cited and citing articles in library and information science journals. *Journal of the American Society for Information Science*, 44(9), 543-552.

He, S. & Spink, S. (2002), A comparison of foreign authorship distribution in JASIST and the Journal of Documentation, *Journal of the American Society for Information Science and Technology*, 53(11), 953–959.

Herrero-Solana, V. & Ríos-Gómez, C. (2006). Producción latinoamericana en biblioteconomía y documentación en el Social Science Citation Index (SSCI) 1966-2003. *Information Research*, 11(2), paper 247. Retrieved February 20, 2014, from http://informationr.net/ir/11-2/paper247.html

Ginn, L.K. (2003). Citation analysis of authored articles in "Library & Information Science Research", 2001-2002. *Mississippi Libraries*, 67(4), 106-109.

Gürdal, O. (2002). Altın yılını yaşayan Türk Kütüphaneciliği dergisinden yansımalar. [Reflections from the Journal of Turkish Librarianship, living its golden year]. *Türk Kütüphaneciliği*, 16(3), 321-328.

Kajberg, L. (1996). A citation analysis of LIS serial literature published in Denmark 1957-1986. *Journal of Documentation*, 52(1), 69-85.

Kim, M.T. (1991). Ranking of journals in library and information science: A comparison of perceptual and citation-based measures. *College & Research Libraries*, 52(1), 24-37.

Kum, İ. & Erdoğan, P.L. (1980). Turkish library developments. *Unesco Journal of Information Science, Librarianship and Archives Administration*, 2(4), 252-256.

Kurbanoğlu, S. (1996a). Türk kütüphanecilerinin kullandığı dergiler. [Journals cited by Turkish librarians]. *Türk Kütüphaneciliği*, 10(2), 105-118.

Kurbanoğlu, S. (1996b). Kütüphanecilik literatüründe atıf hataları. [Citation errors in library literature]. *Türk Kütüphaneciliği*, 10(2), 137-148.

Mukherjee, B. (2009). Journal of the American Society for Information Science and Technology (2000-2007): A bibliometric study. *IFLA Journal*, 35(4), 341-358.

Ramesh, L.S.R.C.V. & Nagaraju, A.V.S.S. (2000). Citation analysis of the Indian Journal of Information, Library and Society. *Indian Journal of Information, Library and Society*, 13(3-4), 171-179.

Schloegl, C. & Stock, W.G. (2004). Impact and relevance of LIS journals: A scientometric analysis of international and German-language LIS journals-citation analysis versus reader survey. J*ournal of the American Society for Information Science & Technology*, 55(13), 1155-1168.

Schubert, A. (2002). The Web of Scientometrics: The statistical overview of the first 50 volumes of the journal. *Scientometrics*, 53(1), 3-20.

Thomson Reuters. (2013). Journal Citation Reports. Retrieved February 20, 2014, from http://admin-apps.isiknowledge.com/JCR/JCR

Tonta, Y. (2002). Türk Kütüphaneciliği dergisi, 1987-2001. [The Journal of Turkish Librarianship, 1987-2001]. *Türk Kütüphaneciliği*, 16(3), 282-320.

Tsay, M. & Shu, Z. (2011). Journal bibliometric analysis: a case study on the Journal of Documentation. *Journal of Documentation*, 67(5), 806-822.

Uzun, A. (2002). Library and information science research in developing countries and Eastern European countries: A brief bibliometric perspective. *International Information & Library Review*, 34(1): 21-33.

Yılmaz, M. (2000). Türkiye'de kütüphane ve enformasyon bilimi alanında bilimsel iletişimin meslek dergilerine yansıması: Bibliyometrik bir analiz. [The reflection of the scientific communication in library and information science field in Turkey to professional journals: A bibliometric analysis]. In A. Yontar (Ed.) *Türkiye'de kütüphane ve enformasyon biliminin kurumsal gelişimi*. (pp. 29-48). İstanbul: Türk Kütüphaneciler Derneği İstanbul Şubesi.

Whitten, B. & Minder, T. (1974). Education for librarianship in developing nations: The Hacettepe experience. *Journal of Education for Librarianship*, 14(4), 220-233.

Comparative Study of Journal Impact Factor and Self-Citation Across Asian International Journals

Wen-Yau Cathy Lin

Associate Professor, Department of Information and Library Science, Tamkang University, Taipei, No.151 Yingzhuan Road, Tamsui District, New Taipei City, 25137, Taiwan
wylin@mail.tku.edu.tw

Abstract

The aim of this research is to explore journal self-citation, as well as its impact on the outcomes of frequently used major citation analysis indicators. Journals published in China, Japan, India, Singapore, South Korea, and Taiwan, during 2010 to 2012 and indexed by JCR-SE are collected and analyzed using citation analysis and statistics methods. The results show that Japan had the highest number of journals indexed in JCR-SE, and also had highest average IF values. Six countries had average self-citation rates mostly below 23%. The self-citation rates of Japan and Singapore are relatively stable, with the lowest values and the least variations. In general, self-citations show no significant effect on corresponding IF values in all countries.

Keywords: self-citation, citation analysis, impact factor, Asian international journal

Introduction

Academic assessment based on citation analysis results has become quite common in recent years. Although it is receiving great attention and popularity, application of citation analysis data on certain evaluation indicators is still the source of much discussion and controversy (Altmann & Gorman, 1999; Bensman, 2012; Lundberg, 2006; Seglen, 1997). In Journal Citations Reports (JCR), there are various indicators based on citation data. These data come from a large number of international journals. Therefore, JCR indicators could represent the journals' global usage (Nisonger, 2000), and they are also useful as a reference for ascertaining the value of academic journals. One of the most controversial issues about citation indicators, such as impact factor (IF) in JCR, concerns the influence and potential manipulation of journal self-citation.

Since 2008, JCR has been simultaneously releasing its new edition each June and announcing the list of suppressed journal titles. In 2008, nine journals were absent from the 2007 edition of JCR because of their exceptionally high self-citation rates. Within the next few years, the suppressed lists became even longer. In the 2011 edition, JCR no longer referred specifically to the excluded journals' high self-citation rates, but rather used “anomalous citation patterns” to indicate the distortion of the journals' impact factors, and the number of suppressed titles climbed to 50. There were 65 suppressed journals in 2012 edition, which is the highest record since 2007.

According to the Science and Engineering Indicators 2014 (National Science Board, 2014), there are five Asian countries, China, Japan, South Korea, India and Taiwan, being included in the top fifteen major producers of the world’s science and engineering articles. Asian countries increasingly show their importance and influence on the international scholarly community. However, during 2010 to 2012, there were journals published in China, Japan, South Korea, and Taiwan being suppressed from JCR-Science Edition (JCR-SE) owing to exceptionally high self-citation rates. One might wonder whether these journals share some common characteristics under the regionalism.

Previous studies about journal performance in JCR focused on certain countries or languages such as China (Zhou & Leydesdorff, 2007), Japan (Zhang & Yamazaki, 1998), non-English speaking countries (González-Alcaide, Valderrama-Zurián, & Aleixandre-Benavent, 2012), Spanish and Swedish (Biglu & Askari, 2007), and French (Bracho-Riquelme, Pescador-Salas, & Reyes-Romero, 1999). There is a lack of multinational comparative studies.

The aim of this research is to explore journal self-citation, as well as its impact on the outcomes of frequently used major citation analysis indicators. Journals published in scientifically and technologically advanced countries in Asia, including China (CN), Japan (JP), India (IN), Singapore (SG), South Korea (SK), and Taiwan (TW), during 2010 to 2012 and indexed by JCR-SE are collected and analyzed using citation analysis and statistics methods. The relationship between journal self-citation rates and citation analysis indicators, such as IF and 5-Year IF, will be investigated, along with changing trends in self-citation rates.

Methods and data collection

The number of Asian international journals collected in the JCR Social Science Edition is significantly fewer than that in JCR-SE. To achieve a comparable data set, this study focuses on academic journals collected only in JCR-SE. Countries with at least 30 journals currently indexed in JCR-SE are selected as research targets. During 2010 to 2012, based on the results of applying the Country/Territory Selection function, journals published in keywords "Japan", "India", "Peoples R China", "Singapore", "South Korea", and "Taiwan" are analyzed. Table 1 indicates the numbers of academic journals included in this research. It is possible that some journals published by academic societies or associations changed their publishing location during the time scope of this research. To avoid any uncertainty, country classification is limited to what is recorded in JCR-SE.

Table 1. Asian International Journals indexed in JCR-SE

	CN	JP	IN	SG	SK	TW	Total
2010	138	207	94	51	75	31	596
2011	155	236	100	50	82	32	655
2012	152	239	105	52	90	33	671

Results and discussion

IF & 5-Year IF

Of these six countries, Japan had the highest number of academic journal titles collected by JCR-SE, followed by China. During 2010 to 2012, South Korea was found to have a significantly increasing number of journal titles in JCR-SE. As for the IF and 5-Year IF comparison (Table 2), the average IF values for journals published within these six countries are all lower than 1.3. There are only journals published in China and Japan had average IF values of higher than 1. However, as revealed by Table 2, their standard deviations are also higher than 1, indicating a wide range of variation across China and Japan journals.

Table 2. Mean and Standard Deviation of IF & 5-Year IF

	2010				2011				2012			
	IF		5YIF		IF		5YIF		IF		5YIF	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
CN	0.96	1.54	1.06	1.72	0.99	1.33	1.15	1.52	1.13	1.40	1.22	1.45
JP	1.09	1.06	1.20	1.08	1.16	1.07	1.29	1.22	1.22	1.07	1.32	1.23
IN	0.47	0.45	0.59	0.46	0.54	0.47	0.64	0.49	0.55	0.48	0.63	0.53
SG	0.82	0.59	0.89	0.49	0.69	0.39	0.71	0.33	0.65	0.34	0.75	0.34
SK	0.78	0.52	0.91	0.57	0.83	0.51	0.94	0.53	0.91	0.54	0.09	0.52
TW	0.77	0.46	0.83	0.44	0.82	0.61	0.93	0.57	0.84	0.59	0.88	0.60

Within this research scope, the journal with the highest IF/5-Year IF value is the *Journal of Molecular Cell Biology* (2010), published in China. This journal belongs to the subject category of *CELL BIOLOGY*. It is an extremely high impact journal and lists in Q1 within a total of 178 journals in this field. In year 2011 and 2012, *Cell Research*, another Chinese journal listed in the same category *CELL BIOLOGY*, had the highest IF value. Regarding journal titles published in Japan, *Journal of Photochemistry and Photobiology C-Photochemistry Reviews*, a Q1 journal in the subject category of *CHEMISTRY, PHYSICAL*, had the highest IF value in both 2010 and 2011. A detailed comparison of the maximum and minimum IF values is given in Table 3.

Table 3. Maximum and Minimum of IF & 5-Year IF

	2010				2011				2012			
	IF		5YIF		IF		5YIF		IF		5YIF	
	Max	Min	Max	Min	Max	Min	Max	Min	Max	Min	Max	Min
CN	13.4	0.058	13.8	0.148	8.19	0.27	7.889	0.027	10.526	0.068	10.216	0.075
JP	10.81	0.03	10.271	0.05	10.36	0.009	12.458	0.021	9.042	0.051	11.952	0.073
IN	2.6	0.035	2.125	0.68	2.722	0.013	2.218	0.123	2.272	0.019	2.306	0.029
SG	3.139	0.1	2.671	0.237	1.781	0.129	1.844	0.253	1.874	0.176	1.788	0.229
SK	2.453	0.091	2.542	0.156	2.481	0.071	2.384	0.174	2.653	0.174	2.513	0.136
TW	1.962	0.085	2.124	0.27	2.827	0.155	2.12	0.216	2.458	0.193	2.339	0.251

In JCR, any particular journal might be classified to a single or multiple subject categories, and the values of IF are ranked separately. This study double counted all of the rank percentage for each journal and calculated the average number of each country annually. Smaller percentage value of category ranking represents a better performance on IF. The rank percentage analysis results show that journals published in China, Japan, and South Korea have values around 60s%, Singapore and Taiwan journals have values close to 70%, and India journals are in the range of 80%. In general, these six Asian counties hold low journal IF value in various categories.

Table 4. IF ranking in Categories

	CN	JP	IN	SG	SK	TW
2010	0.685	0.678	0.835	0.668	0.669	0.713
2011	0.669	0.677	0.824	0.706	0.680	0.723
2012	0.658	0.664	0.819	0.732	0.659	0.728

Self-citation rate

The self-citation rates calculated from this research are listed in Table 5. Among journals published in the six countries under investigation, average self-citation rates are mostly below 23%. Self-citation rates reported from Japan and Singapore are relatively stable, with not only the lowest values but also the least variations. Apart from Singapore, self-citation rates for journals published in the other five countries did decrease with time. However, there are some extreme cases worth addressing. For example, one journal published in India was found to have an extraordinary self-citation rate of 100% in 2011. JCR-SE immediately suppressed this journal in 2012. Another journal, also published in India, had a high self-citation rate of 89.29% in 2010. This value dropped to 70% in 2011 but increased to 83.67% in 2012. It is worthwhile to pay close attention to whether this journal will be included in JCR-SE in the future. In China, one journal had the highest self-citation rate of 82.14% in 2011. This is also the first year that journal was indexed in JCR-SE. In the next year, its value dropped significantly to 20%.

Table 5. Self-citation rate of Asian International journals (%)

	2010				2011				2012			
	Mean	SD	Max.	Min.	Mean	SD	Max.	Min.	Mean	SD	Max.	Min.
CN	20.54	15.85	73.85	1.33	19.02	14.65	82.14	0.55	17.59	14.21	79.77	1.22
JP	10.84	10.43	69.99	0.18	10.15	9.11	64.52	0.48	9.95	9.23	72.30	0.56
IN	16.93	15.71	89.29	0.92	15.91	15.90	100	0.98	15.1	15.52	83.67	0.39
SG	11.45	11.65	58.43	0.96	9.97	7.960	39.60	1.41	10.96	9.91	55.41	1.97
SK	22.74	17.08	75.85	0.66	20.35	16.03	62.50	2.22	19.02	14.68	56.52	1.16
TW	19.57	16.92	72.22	1.6	17.42	16.45	62.85	1.85	16.45	14.42	52.58	1.20

IF & Self-citation

To explore how IF values are affected by self-citation, statistical means (Pearson's) on the IF values calculated between including/excluding self-citations were collected and listed in Table 6. Data from the six countries show significantly correlated levels. This is an indication that even though there are some individual journals with extremely high self-citation rates, in general, self-citations show no significant effect on corresponding IF values. It cannot be excluded that this has contributed to the long-term gatekeeping role played by Thomson Reuters.

Table 6. Correlation Coefficients (Pearson's *r*) between IF including/excluding self-citations

	CN	JP	IN	SG	SK	TW
2010	.987**	.983**	.985**	.808**	.947**	.866**
2011	.987**	.988**	.985**	.985**	.965**	.808**
2012	.992**	.989**	.860**	.952**	.947**	.911**

**significantly correlated when the significance level is set at 0.01 (two-tailed)

Conclusion

This research is to explore journal self-citation rate and its impact on the journal impact factor. Journals published in China, Japan, India, South Korea, and Taiwan, during 2010 to 2012 and indexed by JCR-SE are analyzed. The results show that Japan had the highest number of journals indexed in JCR-SE, and also had highest average IF values. In generally, these six Asian counties hold low journal IF values in various categories. Six countries had average self-citation rates mostly below 23%. The rates of Japan and Singapore are relatively stable, with not only the lowest values but also the least variations. There are some individual journals with extremely high self-citation rates, in general, self-citations show no significant effect on corresponding IF values in all countries.

Acknowledgments

This research is supported by National Science Council of Taiwan under grant NSC102-2410-H-032-086.

Reference

Altmann, K. G., & Gorman, G. E. (1999). Can impact factors substitute for the results of local use studies? Findings from an Australian case study. *Collection Building, 18*(2), 90-94.

Bensman, S. J. (2012). The impact factor: its place in Garfield's thought, in science evaluation, and in library collection management. *Scientometrics, 92*(2), 263-275. doi: 10.1007/s11192-011-0601-9.

Biglu, M. H., & Askari, O. (2007). Comparison of Spanish and Swedish Journal Indicators (Impact Factor and Self-citation Rate) in the Journal Citation Reports. *Libres, 17*(2).

Bracho-Riquelme, R. L., Pescador-Salas, N., & Reyes-Romero, M. A. (1999). The change from French to English and its effect upon the impact factor and ranking of the Pasteur journals. *Journal of Information Science, 25*(5), 413-417. doi: 10.1177/016555159902500507.

González-Alcaide, G., Valderrama-Zurián, J., & Aleixandre-Benavent, R. (2012). The Impact Factor in non-English-speaking countries. *Scientometrics, 92*(2), 297-311. doi: 10.1007/s11192-012-0692-y.

Lundberg, J. (2006). *Bibliometrics as a research assessment tool - impact beyond the impact factor.* Ph. D., Karolinska Institutet, Stockholm, Sweden.

National Science Board (2014). *Science and Engineering Indicators 2014*. Arlington, VA. : National Science Foundation (NSB 14-01).

Nisonger, T. E. (2000). Use of the Journal Citation Reports for serials management in research libraries: an investigation of the effect of self-citation on journal rankings in library and information science and genetics. *College and Research Libraries, 61*(3), 263-275.

Seglen, P. O. (1997). Why the impact factor of journals should not be used for evaluating research. *British Medical Journal, 314*(7079), 498-502.

Zhang, H., & Yamazaki, S. (1998). Citation indicators of Japanese journals. *Journal of the American Society for Information Science, 49*(4), 375-379.

Zhou, P., & Leydesdorff, L. (2007). The citation impacts and citation environments of Chinese journals in mathematics. *Scientometrics, 72*(2), 185-200.

Statistical analysis on interlocking directorate in Chinese listed companies

Xiaoyu Zhu, Zeyuan Liu, Chaomei Chen and Haiyan Hou

WISE Lab, Dalian University of Technology, Dalian, China 116085

Abstract

In this paper, we use statistical analysis methods to analyze the interlocking directorate in Chinese listed companies. Results showed that the number of the interlocking directorate sustain growth from 2003 to 2013. The proportion of the interlocking directorate in board increased year after year. And independent directorate is the most main part in interlocking directorate. At last we analyze the sex and age distribution of interlocking directorate and the board of directors.

Keywords: Interlock director, independent director, board of directors, Chinese listed company

Introduction

Ever since the birth of the modern corporation, interlock relations that tie firms together in networks of ownership and control have been in place. Interlocking directorates is defined by Mizruchi as the situation where a person affiliated with one organization sits on the board of directors of another organization[1]. Interlocking directorates, where two firms share at least one director, in figure 1number 1-4 represent 4 firms and letter A to K represent the board of directors. Directors A to E constituted the board of directors of firm1, directors B to G constituted the board of directors of firm 2 and so on. Then we found director B to E affiliated firm 1 and firm 2 both. We called them interlocking directorates

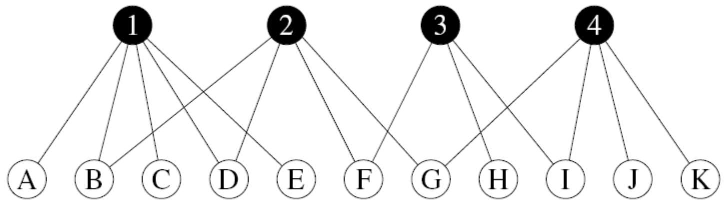


Fig.1 Network structure of interlocking directorate

There are a lot of researches on some counties. Rolfe analyzed interlocking directors in Australia[2].Khanna computed interlock phenomenon in Chile[3].And there had some researches about Italy[4-6]. Keister published the first article about China, but it is just include 40 firms and 535 directors from 1988 to 1990 [7].

Data sources and processing

We got all the 436,285 executives data in CSMAR database from 1999 to 2013, after then we extracted 225,056 directors from the executives. At last, the time span is set as from 2003 to 2013, because the data it is no complete from 1999 to 2002. Another important reason for this is that CSMAR didn't provide directors' resume until 2003. Without the directors' resume, it is very difficult for us to distinguish the director of same name. In China, same first name and last name, it is not just the few.

Results

Number distribution of interlocking directorate

Ever since the birth of interlocking directorate, the number of interlock directors has present a growth trend year by year. From 2003 to 2013, interlock directors present negative growth just 2005 and 2013. There has sustained growth in the rest of the years. And the data showed 22.29% growth in the number of interlock directors in 2010.

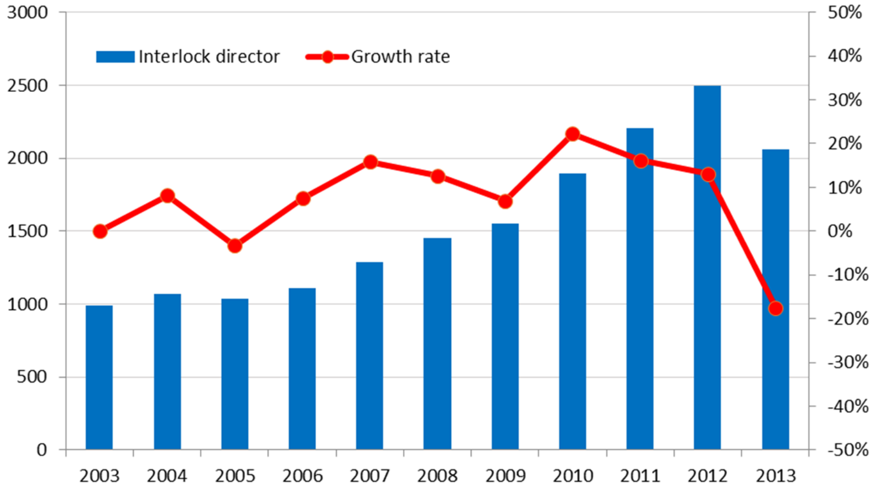


Fig.2 Number of interlocking directorate and growth rate from 2003 to 2013

Interlocking directorate in the proportion of the board

In recent years, the number of interlock directors has present a growth trend, they occupy more and more positions in the board. Figure1shows the number of interlocking directorate and growth rate from 2003 to 2013. Over the past 11 years, the number of interlocking directorate from 2003 to 2013 keeps a relatively stable growth from 988 in 2003 to 2,058 in 2013, with a growth rate of 108%. If remove 2013, the number of interlocking directorate with a growth rate of 152%.

In 2003, all of the Chinese listed companies had 2,991 directors and 988 interlocking directors. The ratio of interlocking directorate in the board is only 33.03%. In 2008, it has increased to 3,459 directors and 1,451 interlocking directors, the ratio increased to 41.95%. In 2011, it has continually increased to 4,216 directors and 2,205 interlocking directors, the ratio increased to 52.3%. The ratio of interlock over half firstly. The peak is 55.34% in 2012.

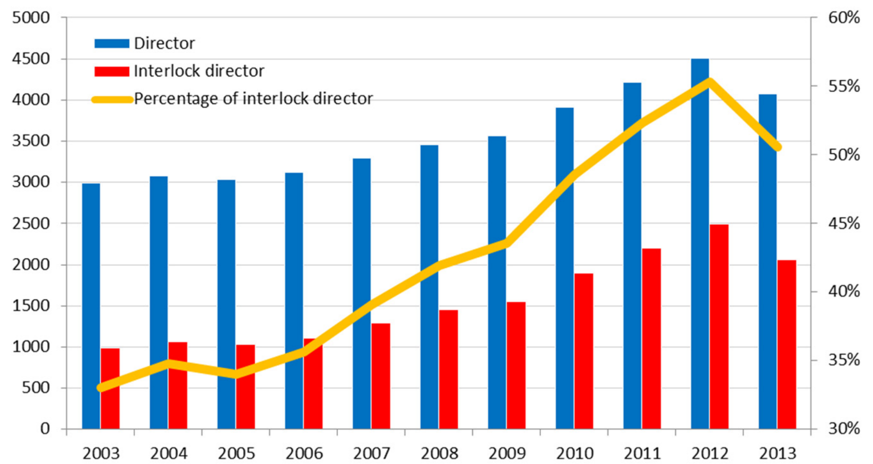


Fig.3 Interlocking directorate in the proportion of the board from 2003 to 2013

Figure 3 illustrates the ratio of interlocking directorate in the board. We can observe from figure 4 clearly, as is shown, from 2003 to 2012, the proportion is in a growth trend overall, the proportion decreased only in 2005 and 2013. Especially, the proportion of interlock over non-interlock for the first time in 2011.

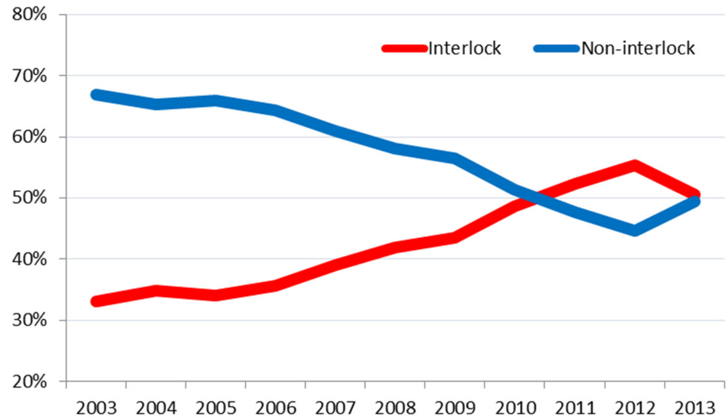


Fig.4 The proportion of interlock and non-interlock 2003 to 2013

Independent director in the proportion of interlocking directorate

Independent director is a very important part in the board of company in China. Independent director is the elite in various industries, especially, college teacher, lawyer and accountant. So it is very necessary to distinguish independent director from interlock director. From 2003 to 2013, with the number of the listed companies accumulated, the number of the board of directors gradually increased, then the number of both independent director and interlocking directors also increased. But from the figure 5, the proportion of independent director in interlocking directorate maintain at 65%-70%. In figure 5, we can compute the proportion of independent directors in interlocking directors. It shows the proportion of 64.16% in 2003, over 70% in 2009; 74.45% in 2013, it's the highest proportion at present. In conclusion, independent directors become a dominant subject in interlocking directors.

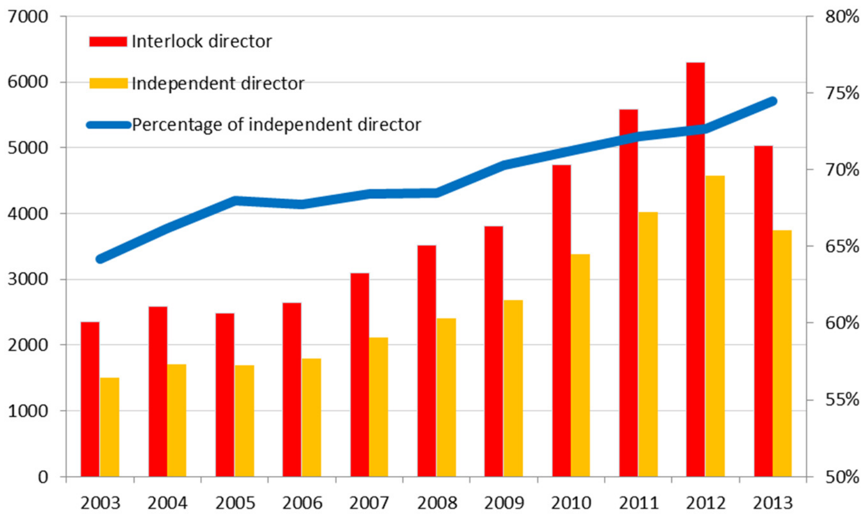


Fig.5 Independent director in the proportion of interlocking directorate

Demography of the interlocking directorate

Sex Distribution Taking “Female director" as an example, there had 10308 male directors and just 1126 female directors in 2003, the proportion of female is only 9.85% ; 16706 male directors and just 2265 female directors in 2013, the proportion of female is only 13.76%. Over the last 11 years, even though the number of female directors increased and the proportion of female also increased, gender difference is still obvious in directors. There also existed gender difference in interlocking directors. There had 907 male directors and just 81 female directors in 2003, the proportion of female is only 8.20% ; 1799male directors and just 259 female directors in 2013, the proportion of female is only 12.59%. The figure 6 shows, whether interlock or not, the number and the proportion of female directors are relatively few. It seems pretty clear that the board of Chinese listed companies still dominated by men in the past 11 years.

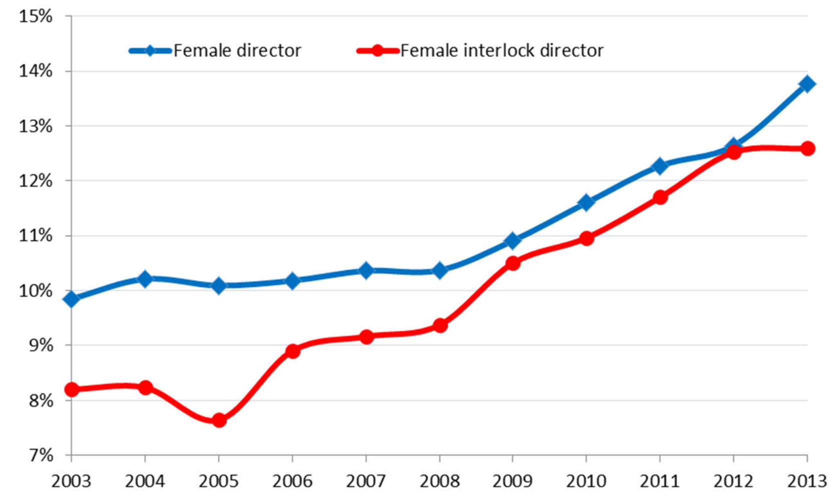


Fig.6 The percentage of female directors and interlock directors

Age distribution Figure 7 shows the age distribution of the board of directors from 2003 to 2013, the age range of directors is from 19 to 102. And figure 8 shows age distribution of interlocking directors from 2003 to 2013, the age range of directors is from 25 to 81. This illustrates that the age of interlocking directors is more centralization. But the most age of the board of directors and interlocking directors are very similar, both around 45 to 55 years old.

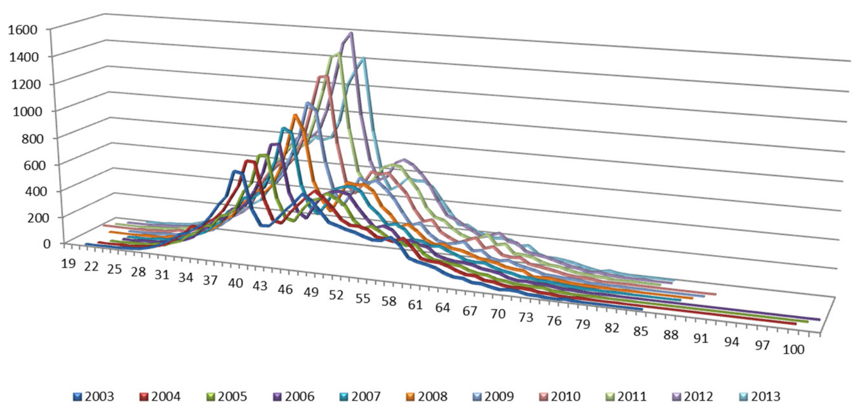


Fig.7 The age distribution of the board of directors

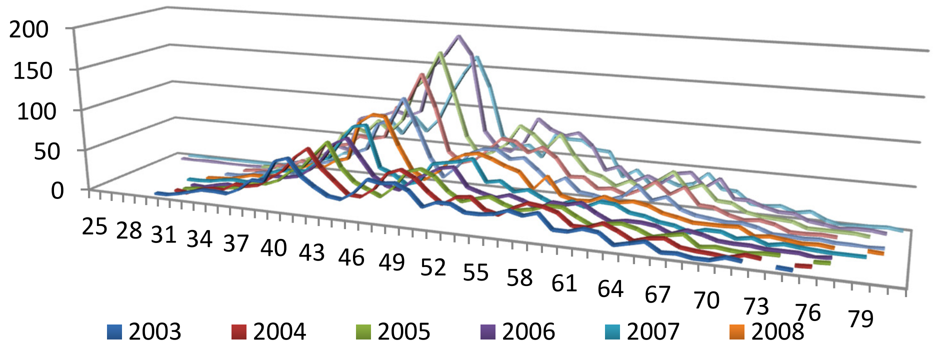


Fig.8 The distribution of the age of interlock directors

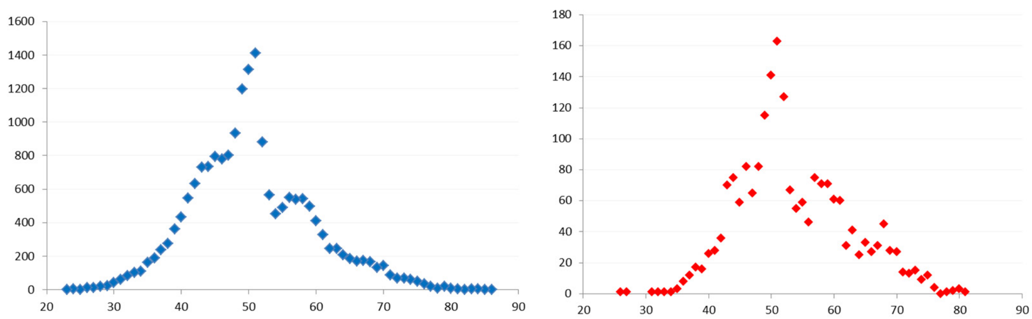


Fig.9 Age distribution of the board of directors and interlock directors in 2013

Figure 9 shows age distribution of directors. In order to observe more intuitive age difference of the board of directors and interlock directors, we put both data together to compare. Because of the number of the board of directors and interlock director, there are considerable differences in data, so we chose logarithmic scale to analyze, see figure 10.

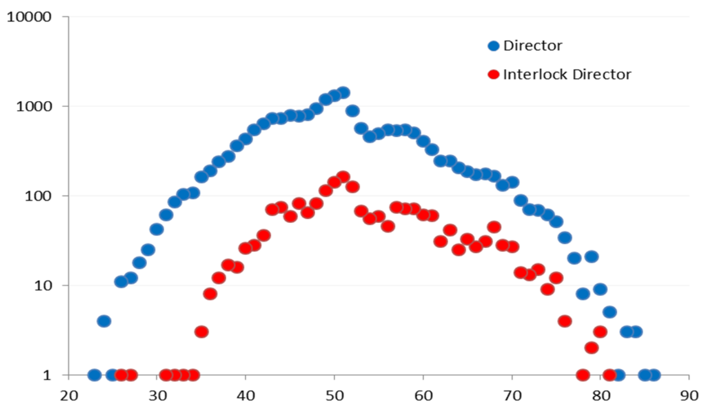


Fig.10 Age distribution comparison in 2013

Conclusion and Discussion

The contribution of this paper is providing a statistics analytical method of calculate interlocking directorate, calculating the number of the interlocking directorate and interlocking directorate in the proportion of the board and independent director in the proportion of interlocking directorate, analyzing sex and age distribution of interlocking directorate and the board of directors. Unlike previous studies focused on a region or industry field, in our work we calculate the interlocking directors of the whole Chinese listed companies for 11 years. The results showed interlocking directorate play such an important role in Chinese listed companies.

This allows us to better quantify further network structure analysis of the interlocking directorate.

Acknowledgments

This work is a periodical research result of the major project of the national social science fund (No. 12&ZD117, 14BTQ030), the project of the natural science fund (71373253, 71103022), the project of the innovation approach (2013IM010100), the project of Chinese degree and graduate education society (2013Y12), the reform project of graduate education of Dalian University of Technology (JG2013002), the planning project of social science of Liaoning province (L13BGL014). Corresponding author: Haiyan Hou, seabirdofsummer@sina.com.

References

[1] Mizruchi M (1996) What do interlocks do? An analysis, critique, and assessment of research on interlocking directorates. Annual review of sociology: 271–298.

[2] Rolfe, H., The controllers: interlocking directorates in large Australian companies. 1967: Cheshire.

[3] Khanna, T. and C. Thomas, Synchronicity and firm interlocks in an emerging market. Journal of Financial Economics, 2009. 92(2): p. 182-204.

[4] Croci, E. and R. Grassi, The economic effect of interlocking directorates in Italy: new evidence using centrality measures. Computational and Mathematical Organization Theory, 2010: p. 1-24.

[5] Vasta, M. and A. Baccini, Banks and industry in Italy, 1911-36: new evidence using the interlocking directorates technique. Financial History Review, 1997. 4: p. 139-160.

[6] Rinaldi, A. and M. Vasta, The structure of Italian capitalism, 1952-1972: new evidence using the interlocking directorates technique. Financial History Review, 2005. 12(2): p. 173.

[7] Keister, L.A., Engineering Growth: Business Group Structure and Firm Performance in China's Transition Economy 1. American Journal of Sociology, 1998. 104(2): p. 404-440.

Impact of Funding Agencies' Activity on Russian Higher Education Sector – The Russian Foundation for Basic Research and foreign funding agencies' collaboration

V.A. Markusova*, A.N. Libkind*, L.E. Mindeli** and E. Noyons***

*Head of Department of the All Russian Institute for Scientific and Technical Information of the RAS, Usievicha 20, Moscow 125190, Russia
markusova@viniti.ru, libkind@viniti.ru

**Director of the Institute for the Study of Science of the RAS, Butlerova 12, Moscow 117485, Russia
l.mindeli@issras.ru

***Deputy Director, Centre for Science and Technology Studies, the Netherlands
Wassenaarseweg 62A, 2333AL Leiden Zuid Holland
noyons@cwts.leidenuniv.nl

Abstract

Competitive funding was introduced in Russia in 1992 with the government's creation of the Russian Foundation for Basic Research (RFBR). Reform of two main Russian research sectors, the Russian Academy of Sciences and the Higher Education Sector (HES) has been going on for the last ten years with the government shifting its attention and financial resources toward the higher education sector (HES). The goal of our empirical project was to give an overview of various funding agencies' (FA) activities supporting the HES; to identify leading universities by number of publications and level of research supported by FA; to examine universities' publications supported only by foreign FA and their subject category's priorities. We performed bibliometric analysis of Russian research output (86,700 records) indexed by Web of Science for period 2009-2011. Research output (RO) supported by various FA for total Russia and RO by the HES were 42,916 records and 18,495 records respectively. About 25% (357) of Russian universities received competitive funding from domestic and foreign FA. It was observed 24 % growth of funded publications numbers between 2009 and 2011. The list of the top 20 domestic and foreign FA supported papers (no than fewer 300) has shown that the RFBR was the leader by number of publications, followed by the Ministry of Education and Science (MES), and programs funded by the Presidium of the Russian Academy of Sciences.

The study revealed an extensive collaborative network of Russians universities with 606 foreign FA contributed to basic research in 160 Russian universities. About 10.6% of analyzed publications were supported only by foreign FA with disciplinary priorities focused on "hard sciences".

Our data indicate that there is a good correlation by Spearman between the share of papers funded by foreign FA and mean-weighted impact factors (MWIF) of these universities' publications (r=+0.78). Despite a very substantial difference in RO of the Moscow State University and St. Petersburg State University compared with other universities, the highest value of mean weighted impact factor and research level were demonstrated by the Moscow Physics Engineering Institute-the National Nuclear Research University.

Our data demonstrate the impact of competitive funding on the Higher Education Sector research activity and provide a better empirical basis for science policy[1]

Keywords: Funding agency, Russia, university, research output, citation score, impact factor, mean-weighted impact factor, international collaboration

Introduction

Russian science policy and Russian bibliometric performance were the subjects of many papers (Wilson, 2004, Lewison & Markusova 2010, Markusova, 2013). Prior to the collapse of the

[1] This paper is a part of the projects supported by the Russian Humanities Foundation (RHF), Grant № 12-00070 and Grant № 14-03-00333.

Soviet Union in 1991, the State funding was the only channel of basic research funding in the country. Competitive funding was introduced in Russia in 1992 with the government's creation of the Russian Foundation for Basic Research (RFBR). At the same time, the opportunity emerged to apply for grants awarded by foreign organizations. Reform of two main Russian research sectors, the Russian Academy of Sciences and the Higher Education Sector (HES) has been going on for the last ten years with the government shifting its attention and financial resources toward the HES. We can state that bibliometric performance of the RAS and the HES played a very important role in this reform. According to a decree released by Prime Minister D. Medvedev (N 979, Nov.1, 2013) http://www.ras.ru/news/shownews.aspx?id=613a30f8-1475-4d9a-a6a3-75df1501be7a, it is mandatory for any research organization's evaluation to include number of papers, their citation scores, and the journal's impact factor indexed by Web of Knowledge or Scopus. The number of grants awarded to an organization is estimated as an indicator of economic performance. We want to emphasize that Russian government science policy is directed towards encouragement of competitive funding. To implement this policy on Oct.16, 2012 it was set up the Foundation of Perspective Research (N174-FZ) with the budget for the period 2014-2016 about $108 mln, $94.3 mln and $106 mln respectively. http://ria.ru/science/20140211/994340988.html#ixzz2w1KO1tF5.

A year later, on Nov.2, 2013 it was set up the Russian Scientific Foundation. Its budget should be 47 bln rub. (about $1.34 bln) for the period 2014-2016. The first call for the proposals was announced on March 1.2014. We believe that more information about criteria' selection of this foundation will be available soon.

The goal of our empirical project was to give an overview of various funding agencies' (FA) activities supporting the HES; to identify leading universities by number of publications and level of research supported by FA; to examine universities' publications supported only by foreign FA and their subject category's priorities; to trace collaboration between RFBR and foreign funding agencies. This paper is follow up study partly discussed in the paper (Markusova, 2013).

Methods

The data for this study have been derived from Thomson Scientific resources: Science Citation Index-Expanded (SCI-E) from Web of Science (WoS) and Journal Citation Reports-Science Edition (JCR) - 2011. All research documents (article, letter, note, and review) with at least one Russian address and indexed between 2009-2011 were downloaded with Thomson Scientific permission. Papers were assigned to a country and Russian institutes based on the address which appears in a paper. Five percent of the records were excluded from analysis due to lack of data.

A more than 88,000 bibliographic records were downloaded from the SCI-E (AD=Russia and PY= 2009-2011)[2]. About 18,500 records contained the information about FA support of the Russian Higher Education Sector (HES). FA names were verified by special software and then checked manually. The result of verification was a list that contained 1,090 FA names or organizations.

Bibliometric indicators: research output (RO) and its share supported by funding agency; RO distribution by university, subject category (SC), country; citation per paper; impact factor (IF); mean-weighted IF (MWIF); aggregated IF (AIF) of subject category (SC.; and_ratio between citations per a funded paper and citations per a paper from total university RO. The file of funding publications is on three levels: funding agency, country, and university.

[2] Search was done on March 10, 2013. Additional search to collect citations was performed on Dec.20, 2013.

To estimate impact of competitive funding on citedness of a university's publications we calculated the ratio between an average citation per a funded paper and an average citation per a paper of total RO by this university. If the ratio is higher than one, it means that it is a positive impact of competitive funding and visa-versa. The file of funding publications is examined on three levels: funding agency, country, and university.

Results and Discussion

Total Russian RO for period 2009-2011 consisted of 86,737 records. The share of papers published by the HES was 43.7%. RO supported by various FA for total Russia RO and RO by the HES were 42,916 records and 18,495 records respectively. Among 1.500 Russian universities, 467 universities contributed papers to WoS, and among them publications from 377 universities were supported by FA.

Table 1. Bibliometric statistics of Russian papers for 2009-2011.

Research output of:	**2009**	**2010**	**2011**
Total Russia	29,097	27,945	29,689
Higher Education Sector (HES)	12,433	12,122	13,447
Share of HES in total Russia RO (%)	42.7	43.4	45.3
HES published in foreign journals	5,221	5,262	5,956
Share of RO published in foreign journals (%)	42.0	43.4	44.3
HES RO supported by FA	5,546	6,073	6,876
Share of RO HES supported by FA	44.6	50.1	51.1
HES supported by Russian FA	4,873	5,362	6,202
Share of HES RO supported by Russian FA	87.9	88.3	90.2
HES RO supported by foreign FA	2,069	2,153	2,167
Share of HES RO supported by Foreign FA (from total HES RO - %)	16.6	17.8	16.1
Share of HES supported by foreign FA (from HES RO supported by all FA - %)	37.3	40.2	34.0
HES RO supported by foreign FA and published in foreign journals	1,627	1,766	1851
Share of HES supported by foreign FA and published in foreign journals	78.6	82.0	85.4

A more 24% growth of publications' numbers supported by FA was observed between 2009 and 2011. It was observed a growth (more than 27%) of publications supported by 119 Russian FA between 2009 and 2011. About 37% of papers were supported by foreign FA, mainly in collaboration with Russian FA.

The share of papers published in foreign journals by total HES RO increased for from 42.0% up to 44.3% in 2009 and 2011 respectively. It is obvious that authors whose publications were supported by foreign FA have been published mainly in foreign journals. Share of these publications was approximately twice higher (above 80%) than the share of papers (about 42%) published in foreign journals for total HES RO.

It was revealed that an extensive network of Russian universities with domestic and foreign FA was operating in 79 countries. Over the past 20 years, two countries, U.S. and Germany, have been the leaders in international collaborations with Russian researchers. Both countries were also leaders in support of competitive funding. An average citation score per paper (10.0) supported by a foreign country for 2009-2011[3] (at the end of 2013) is a little higher than average citation score per Russian internationally collaborated paper in total HES RO (8.2). However, it is in 3.5 folds as high as an average citation per Russian paper for 2008-2012 according to InCites (2.75).

Bibliometric characteristics of top 20 domestic and foreign FA (supported no than fewer 300) is presented in Table 2. Agencies are ranked by RO.

Table 2. Bibliometric characteristics of top 20 funding agencies supported the HES.

Columns: 1 – FA name; 2 – Country; 3 – Funded HES RO; 4 – Share of papers (from total HES RO - %); 5 – Mean-weighted impact factor (MWIF); 6 – Citations per a funded publication; 7 – Ratio between an average citations per a funded paper and an average citations per a paper of total RO

N	1	2	3	4	5	6	7
1	RFBR	Russia	12,875	69.6	1.45	3.8	1.13
2	Ministry of Education and Science (MES) of Russian Federation (RF)	Russia	3,863	20.9	1.81	5.1	1.53
3	Russian Academy of Sciences (RAS)	Russia	2,432	13.1	1.63	4.9	1.46
4	Grants of President RF for Support of Leading Scientific Schools	Russia	1,917	10.4	1.18	3.0	0.89
5	German Research Foundation	Germany	1,174	6.3	3.69	12.1	3.60
6	Federal Targeted Program “Scientific and Scientific-Pedagogical Personnel of the Innovative Russia”	Russia	927	5.0	0.98	1.9	0.56
7	US National Science Foundation (NSF)	US	833	4.5	4.69	20.4	6.06
8	Grants of President RF for Support of Young Scientists; Russia	Russia	729	3.9	1.65	4.1	1.21
9	US Department of Energy (DOE);	US	623	3.4	4.98	21.3	6.32
10	US Civilian Research and Development Foundation (CRDF)	US	456	2.5	1.77	5.7	1.69
11	French National Center for Scientific Research (CNRS)	France	450	2.4	4.44	21.1	6.25
12	National Natural Science Foundation of China	China	438	2.4	4.37	18.6	5.50
13	European Commission	EU	423	2.3	3.39	10.7	3.16
14	“Federal Target program”	Russia	356	1.9	1.27	2.4	0.73
15	Science and Technology Facilities Council (STFC)	UK	354	1.9	5.08	25.4	7.53
16	National Research Foundation of Korea (NRF)	Korea	321	1.7	5.10	23.9	7.10
17	Natural Sciences and Engineering Research Council of Canada (NSERC)	Canada	316	1.7	4.41	18.9	5.61
18	German Federal Ministry of Education and Research (BMBF)	Germany	312	1.7	4.68	21.8	6.47
19	Zimin Dynasty Foundation	Russia	311	1.7	2.57	5.9	1.75
20	Foundation for Fundamental Research on Matter (FOM)	Netherlands	309	1.7	5.04	21.9	6.49

[3]To collect citations data, the search was performed in December, 2013.

We can see, the Russian Foundation for Basic Research (RFBR) was the leader by number of publications, followed by the Ministry of Education and Science (MES), and programs funded by the Presidium of the RAS. To our surprise, the Federal Target Program (FTP) "Scientific and Scientific-Pedagogical Personnel of the Innovative Russia" of the MES have resulted in less RO (927) than RO (1174) supported by German Research Foundation (DFG). We want to emphasize that there is a large discrepancy in the average amount of money per grant per year among the RFBR, the Russian Humanities Foundation (RHF), and the FTP of the MES: $20,000, $10,000, and $40,000, respectively.

The highest value MWIF (5.10) had papers funded by National Research Foundation of Korea, the Science and Technology Facilities Council of UK (5.08), Foundation for Fundamental Research on Matter (Netherlands - 5.04), US Department of Energy (4.98), US National Science Foundation (4.69), German Federal Ministry of Education and Research (4.68) and CNRS (France - 4.44). These organizations have been distinguished by a large number of citations per an average paper (20 and more). Among top 20 FA there are seventeen whose support had a significant positive impact of citations of a funded paper (see column 7). Unfortunately three very popular government programs had a negative impact on citations of funded paper. These programs are: the Grants of President RF for Support of Leading Scientific Schools (ratio=0.89), the Federal Targeted Program 'Scientific and Scientific-Pedagogical Personnel of the Innovative Russia (0.56) and the "Federal Target program" (0.73).

One of the main goals of our project was to identify leading universities by bibliometrics indicators. We identified 15 universities spread across the vast territory of Russia that have published not less than 150 papers supported by FA. This list includes four Federal Universities and nine National Research Universities (NRU), that are flagmen of the Russian higher education sector. Table 3 contains bibliometric indicators of 15 leading universities

Table 3. Bibliometric indicators of 15 leading Russian universities.

Columns: 1 – University's name; 2 – RO of university funded by FA; 3 – Share of funded RO (%); 4 – Citations share of funded RO (%); 5 – Number of citations per a paper of total university RO; 6 – Number of citations per a funded paper; 7 – Mean weighted impact factor (MWIF) of total university RO; 8 – MWIF of funded RO

University's name	2	3	4	5	6	7	8
M.V. Lomonosov Moscow State University	6057	61,8	75,9	4,6	5,7	1,8	1,9
Saint Petersburg State University	1637	59,4	74,7	5,0	6,3	1,8	2,1
Novosibirsk State University	1085	69,3	73,9	3,9	4,1	1,8	1,9
Moscow Institute of Physics and Technology	582	66,6	76,5	3,6	4,2	1,7	1,9
B.N.Yeltsin Ural Federal University	477	54,5	70,5	2,8	3,7	1,2	1,4
Kazan (Volga Region) Federal University	460	59,7	75,1	3,7	4,7	1,7	2,0
N.I. Lobachevsky State University of Nizhniy Novgorod	437	67,5	75,2	2,3	2,6	1,2	1,3
Southern Federal University	435	52,3	68,5	2,7	3,5	1,3	1,5
Moscow Engineering Physics Institute (MEPhI)	421	47,7	77,4	8,9	14,5	2,0	2,8
Tomsk State University	367	59,2	68,1	2,3	2,6	1,0	1,2
Siberian Federal University	317	60,0	76,8	3,2	4,0	1,3	1,5
Saint Petersburg State Polytechnical University	309	48,7	72,3	4,3	6,4	1,6	2,2

N.G. Chernyshevsky Saratov State University	302	58,2	85,1	4,7	6,9	1,5	1,8
Voronezh State University	250	50,3	70,2	2,2	3,1	0,9	1,2
Tomsk Polytechnic University	243	51,1	72,1	2,8	3,9	1,1	1,4

Leading universities demonstrated a higher share of citations than share of funded papers. We want to emphasize that these universities citations shares for three years period (2009-2011) are significantly higher than citation shares of total Russian RO for 2008-2012 by InCites (48.04%). The Moscow State University and St. Petersburg State University occupy a special position in HES. As a consequence, there is a significant discrepancy in their total RO compared with the RO of other universities. However, by value of MWIF and Research level (RL) the first rank belongs to the Moscow Physics Engineering Institute - the National Nuclear Research University. MWIF of funded publications is slightly higher than MWIF of total university's RO.

To estimate impact of competitive funding on quality and quantity of universities, we randomly selected 85 universities located in 37 cities and 34 regions, which published at least 50 papers in WoS for the studied period. We discovered using Spearman correlation (r) that there is a significant correlation between share of papers funded by foreign FA and the MWIF of these universities' publications (r=+0.78). It was observed that is relatively strong influence of share of all funded papers on total university RO (r=+0.51). We found out a weak positive correlation (r=+0.006) between the share of teachers with a scientific degree and the MWIF of papers funded by all FA. The correlation between share of teachers with a scientific degree and MWIF of papers funded by foreign FA was a little bit higher (r=+0.025). Nevertheless we could assume that scientific degree does not have influence on teachers' choice to publish results in high impact journals.

RO funded by Russian FA was assigned to 174 subject categories (SC) (among 176) of WoS. Among the top 10 SC (each containing 600 or more papers) five were related to "Physics and Astronomy," (in particular "Physics, Multidisciplinary", "Astronomy & Astrophysics", "Physics, Applied", "Physics, Condensed Matter"), two - to "Chemistry" one - to "Mathematics," one - to "biochemistry and molecular biology," and one - to "Optics." An average citation per paper varies strongly by field of science. SC "Biochemistry and Molecular Biology" has very high citation per paper. In the Russian case it is different. SC "Astronomy and Astrophysics" has the highest citation scores per paper (7.0) than SC "Biochemistry and Molecular Biology" (5.0). The highest value of MWIF has SC "astronomy and astrophysics" (2.7) and "Biochemistry and Molecular Biology" (2.2).

Our data revealed that the publications supported by the RFBR were published with the collaboration of 577 foreign FA. We selected fifteen foreign agencies that supported no fewer than 150 papers. Bibliometric characteristics of these leading agencies is presented in Table 4.

Table 4. Top 15 foreign FA collaborating with RFBR, 2009 - 2011, SCI-E.

1 – FA name; 2 – Country; 3 – RO of foreign FA; 4 – Share of RO (from total RO supported by RFBR - %); 5 – Share of citation of this RO (%); 6 – MWIF of RO 7 – Average number citations per paper; 8 – Ratio between average number of citation per funded paper to average number of citation per paper in total RO by RFBR

1	2	3	4	5	6	7	8
German Research Foundation (DFG);	Germany	765	5.9	16.3	3.2	10.4	2.7
US National Science Foundation (NSF);	US	409	3.2	12.2	4.0	14.6	3.8
US Civilian Research and Development Foundation (CRDF);	US	378	2.9	3.9	1.6	5.0	1.3
French National Center for Scientific Research (CNRS);	France	266	2.1	8.6	3.9	15.9	4.2

International Association for the Promotion of Cooperation with Scientists from the New Independent States of the Former Soviet Union (INTAS);	EU	244	1.9	2.6	1.5	5.1	1.3
US Department of Energy (DOE);	US	243	1.9	9.6	4.7	19.2	5.1
National Natural Science Foundation of China (NNSFC, NSFC);	China	237	1.8	8.3	4.2	17.2	4.5
Natural Sciences and Engineering Research Council of Canada (NSERC);	Canada	185	1.4	7.1	4.5	18.6	4.9
Swedish Research Council (SRC)	Sweden	174	1.4	9.7	5.2	27.2	7.2
German Federal Ministry of Education and Research (BMBF)	Germany	169	1.3	7.0	4.5	20.3	5.3
Science and Technology Facilities Council (STFC);	UK	161	1.3	7.8	5.5	23.7	6.2
National Research Foundation of Korea	Korea	157	1.2	6.8	5.2	21.2	5.6
Foundation for Fundamental Research on Matter (FOM);	Netherlands	156	1.2	6.5	5.0	20.2	5.3
National Council of Scientific and Technological Development	Brazil	154	1.2	6.2	4.9	19.7	5.2
European Commission	EU	154	1.2	3.2	2.7	10.2	2.7

As we can see each of this agency has significantly higher citation share than share of funded papers. An average impact factor of Russian journals is very low (0.5). As we can see (column 6), MWIF of these collaboratively funded papers much higher. We could assume that these papers have been published in high impact foreign journals. From our point of view, authors choice is related to selection criteria process. Review criteria of foreign foundations do not mention what role the journal's impact factor (IF) plays. However, by NSF guidance, a reviewer has to evaluate "How well qualified is the proposer (individual or team) to conduct the project? (If appropriate, the reviewer will comment on the quality of prior work)" http://www.nsf.gov/pubs/gpg/nsf04_23/3.jsp.

We assume that one of additional criteria could be the list of principal investigator (PI) publications. One of the authors contacted her Russian colleagues who work abroad and serve as a peer –reviewer for foreign FA. According to e-mail message of molecular biologist, who serves as reviewer in the NSF, U.S. "there is no indication from NSF to use impact factor as an indicator of quality. Nevertheless, any reviewer knows that a publication in "Science", "Cell" or "Nature" is the indicator of high quality research". Another Russian colleague who is senior researcher (immunologist) at the Leicester University, and serves as a peer reviewer for various British FA, noted that a value of IF plays a significant role. If PI does not have a publication in "Cell" (IF=31.957 by JCR 2012) but have been published in journals with IF between 7-10, it is impossible to get the substantial grant (about one mln pounds). Other explanation of this correlation could be a scope of project. Our data show that paper supported by foreign FA has usually a few sponsors and a significant research team. Taking into consideration linguistic barrier it is obvious that foreign partners facilitate a Russian researcher's publication in foreign journal with high impact factor.

A special interest for us was the file of 1,960 publications funded only by foreign FA. Its analysis allows us to identify disciplines, which attract foreign investment in Russian basic research. 606 foreign FA, located in 68 countries contributed to basic research in 183 Russian universities. The top ten countries ranked by number of funded papers are presented in Table

5. As we can see, an average citation per paper is very high compared to an average citation per paper in total Russia RO for 2008-1012 (2.1) by InCites.

Table 5. Top ten countries ranked by RO supported by foreign FA

Rank by RO	Country	RO	Citations/ per a paper	MWIF
1	US	574	18.0	4.37
2	Germany	436	14.1	3.97
3	EU	378	12.2	3.62
4	UK	163	19.4	4.00
5	Spain	156	19.9	4.03
6	France	126	16.5	4.57
7	International	107	4.6	1.58
8	China	89	13.5	3.08
9	Italy	84	24.2	4.74
10	Switzerland	70	23.1	4.66

If we look at the foreign FA distribution by number of papers we get a different picture. The leading foreign FA was the German Research Foundation (224 papers) followed by the NSF USA (189 papers), European Commission (179), and NIH USA (115 papers).

The analyzed file was assigned to 155 SC by WoS. Among the top 13 SC by number of papers nine SC belong to various sub-fields of physics: one SC to "nanotechnology." four SC to chemistry, and only one to "biochemistry and molecular biology.”

Visualization of subject priorities by three foreign FA and one Russian was created using software VOSviewer http://www.vosviewer.com and presented at Fig.1. Cluster 1 belongs to German Research Foundation (DFG); cluster 2 to British Science and Technology Facilities Council (STFC); cluster 3 to the National Institutes of Health (NIH); and cluster 4 to private Russian foundation – Vladimir Zimin “Dynasty” Foundation. The highest number of publication in SC was 247. As we can see, three FA are heavily focused on "hard sciences" and NIH on life sciences.

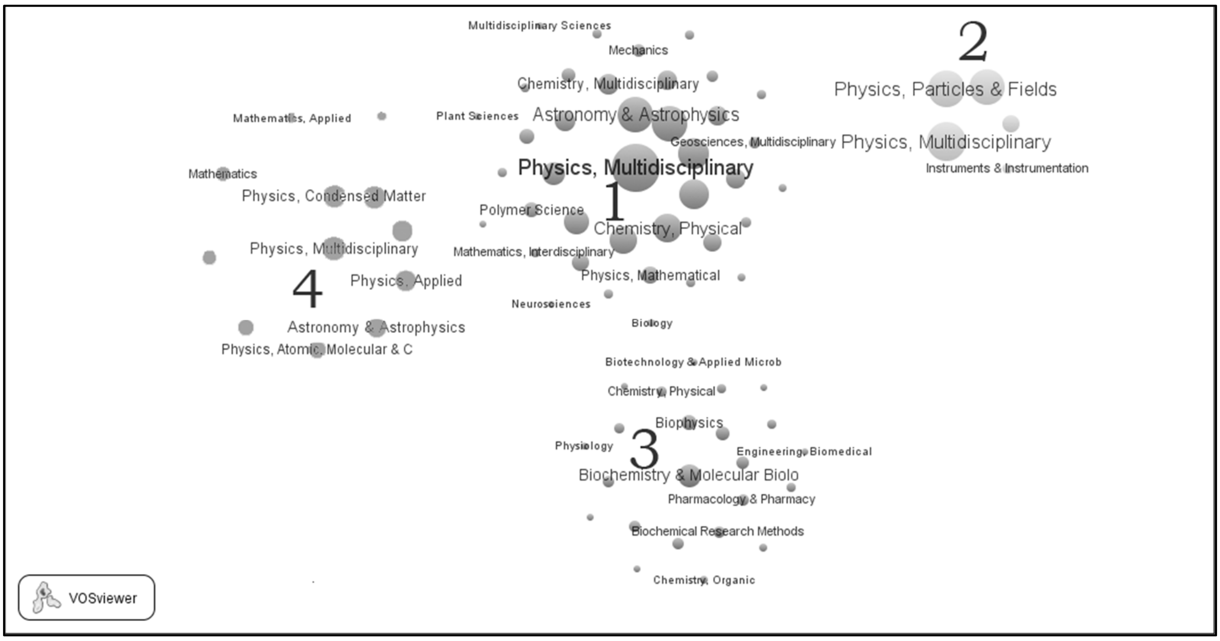


Figure1. Subject’s priorities of various funding agencies.

Conclusions

Short history of government science policy towards competitive funding has proved its positive impact of Russian research community. About 25% (377) of Russian universities received competitive funding from domestic and foreign funding agencies in 2009-2011. Analysis of their research output (18,497 papers) has demonstrated almost 24% growth between 2009 and 2011. The study revealed an extensive collaborative network of Russians universities with foreign FA. About 10.6% of analyzed publications were supported only by foreign FA with disciplinary priorities focused on "hard sciences".

Our data indicate that there is a good correlation by Spearman between the share of papers funded by foreign FA and mean-weighted impact factors (MWIF) of these universities' papers ($r = +0.781$). Despite a very substantial difference in RO of the Moscow State University and St. Petersburg State University compared with other universities, the highest value of MWIF and research level were demonstrated by the Moscow Physics Engineering Institute-the National Nuclear Research University. It was observed a weak correlation between the share of teachers with a scientific degree and value MWIF.

Bibliometrics has become a very important tool in Russian government science policy. Our data demonstrate the impact of competitive funding on the Higher Education Sector research activity and provide a better empirical basis for science policy.

References

Cronin, B., McKenzie, G. & Rubio, L. (1993) The Norms of Acknowledgement in four Humanities and Social Sciences Disciplines. *Journal of Documentation*, 49, 29-43.

Lewison, G., Jeshin, D. & Anderson J. (1995) A Bibliometric Database for Tracking Acknowledgements of Research Funding, *Proceedings of the Fifth Biennial Conference of the International Society for Scientometrics and Informetrics* (pp.255-264). Rosary College, River Forest, IL, USA.

Lewison, G. & Markusova, V. (2010) The Evaluation of Russian Cancer Research. *Research Evaluation*, 19, 2, 129-144.

Markusova, V.A. (2001) Communication Network of Russian Researchers: Acknowledgement Analysis of Russian Journals "Biochemistry" and "Geochemistry". *Scientific and Technical Information*, part.1, 9, 21-28.

Markusova V.A. Research performance by Federal and National Research Universities and impact of competitive funding on their publication activity. *Collnet Journal of Scientometrics and Information Management*, - 2013 - 2(7), - P. 1-11.

Tiew, W.S. & Sen B.K. (1999) Acknowledgement Patterns in Research Articles: A Bibliometric Study Based on Journal of Natural Rubber Research 1986-1997, *Proceedings of the Seventh Conference of the International Society for Scientometrics and Informetrics (*pp.488-497). Mexico: Colima University.

Wang, J. & Shapira, P. (2011) Funding Acknowledgement Analysis: an Enhanced Tool to Investigate Research Sponsorship Impacts: the Case of Nanotechnology. *Scientometrics*, 87, 563-569. (DOI 10.1007/s11192-011-0362-5).

Wilson, C. & Markusova, V.A. (2004) Changes in the Scientific Output of Russia from 1980-2000, as Reflected in the Science Citation Index in Relation to National Politico-economic Factors. *Scientometrics*, 59, 3, 345-389.

Temporal Analysis on Pairs of Classified Index Terms of Literature Databases

Zheng Ma* and Karsten Weihe**

*Information Center for Education, German Institute for Educational Research (DIPF), Schloßstraße 29, 60486 Frankfurt, Germany
ma@kdsl.informatik.tu-darmstadt.de

** Computer Science Department, Technische Universität Darmstadt, Hochschulstrasse 10, 64283 Darmstadt, Germany
weihe@cs.tu-darmstadt.de

Abstract

Trend analysis and anomaly detection (Chan & Mahoney, 2005; Wei, Kumar, Lolla, & Keogh, 2005) is gaining more and more interest since more than a decade. It gains focus in the new era of web2.0. As part of the project *Knowledge Discovery in Scientific Literature*, we developed new promising perspectives to detect trend and other interesting temporal patterns of index terms in the literature databases of educational domain. More specifically, we assign categories to index terms and investigate the index term pairs of special interest. We designed several measures to capture the characteristics of the evolution of individual index terms and pairs of index terms. Result shows our methodology is effective to find interesting temporal patterns, e.g. dependency relationship between index terms and helpful to detect trend.

Introduction

In the last two decades (Allan, Carbonell, & Doddington, 1998), trend detection has become a focus of research. To detect trend in time series, researchers applied various approaches. In the beginning, many systems like TOAK (Porter & Detampel, 1995), CIMEL (Blank & Pottenger, 2001) and ThemeRiver (Havre, Hetzler, & Nowell, 2000) are designed as interactive tools providing statistics and visualization that help user to effectively detect trends. Another type of systems implement some learning algorithm to automatically detect trends. TimeMines (Swan & Jensen, 2000), PatentMiner (Lent, Agrawal, & Srikant, 1997) and HDDI (Pottenger, Kim, & Meling, 2001) are the first ones in this category. With ever growing interest of research, this field has been broadened in two dimensions. Firstly, the target of detection is generalized from trend to interesting patterns (Chan & Mahoney, 2005; Wei, Kumar, Lolla, & Keogh, 2005). Secondly, research in time series data mining greatly enriches the methodology available. The well-know CiteSpaceII (Chen, 2006) represents the successful trend detection system that uses citation data.

In this paper, we present our work in interesting temporal pattern detection in scientific literature database. It provides a means to monitor a scientific field, which is meaningful for researchers as well as decision makers. We designed a series of statistical metrics to find interesting temporal patterns of different types. To the best of our knowledge, we are also the first to employ systematic category information of index terms in such tasks. Our hypothesis is that with category information, it is easier to obtain more interesting and insightful index term pairs.

We use the PEDOCS (DIPF) dataset for experiment. At the moment, it contains more than 5,000 research articles in education domain with full text and rich metadata.

This paper is organized as following. *Related work* introduces relevant important works. *Methodology* present our methods in depth. *Empirical results* summarizes the main findings from our experiments. The last section is *Conclusion and Outlook*.

Related work

Kontostathis et al. (2004) provides a survey of early fundamental work in the trend detection field. In the following, we focus on the work that is most closely related to ours.

The TOAK system (Porter & Detampel, 1995) uses co-occurrence of index terms for visualization, but no calculation and automatic analysis is involved.

To analyse the time series, works have been done in segmentation and similarity measurement of time series. Keogh (1997) uses piecewise linear representation to approximately represent and segment time series. It is similar to our *linear regression segmentation*. However, we compute the exact optimal approximation and segmentation (up to four segments) in comparison to (Keogh, 1997), which uses heuristic methods.

To generalize and smoothen the time series, there is another family of time series representation methods. Agrawal et al. (1995) introduces a shape definition language (SDL) to describe different types of curve pieces, e.g. peak, valley, increasing or decreasing. Further work in that direction (Aref, Elfeky, & Elmagarmid, 2004; Motoyoshi, Miura, & Watanabe, 2002; Yang & Zhao, 1998) convert the segmented time series into predefined symbols. In contrast, pattern detection is not limited to predefined shapes in our work.

TimeMines (Swan & Jensen, 2000) uses a default model as a baseline to detect anomalies. This idea is similar to our Deviation from Random, by which we define the random co-occurrence as null model.

Methodology

Linear Regression Segmentation

As a fundamental analysis of the rough evolution of a single index term, we developed the linear regression segmentation. It is also used in the term pair analysis.

In this paper, we use the term frequency curve of an index term. It refers to the temporal data per index term, for each year (from 1980 to 2013) the number of publications that contain this index term in specific part (full text, abstract or index term list).

Since the frequency curve can be very noisy as can be seen in Figure 1, it is quite hard to automatically summarize the phases in the development of a term frequency. We used Linear Regression Segmentation to segment the frequency curve between 1980 and 2013. The algorithm of Linear Regression Segmentation computes the exact optimal segmentation that best balances the number of segments against the overall approximation error. It enumerates all

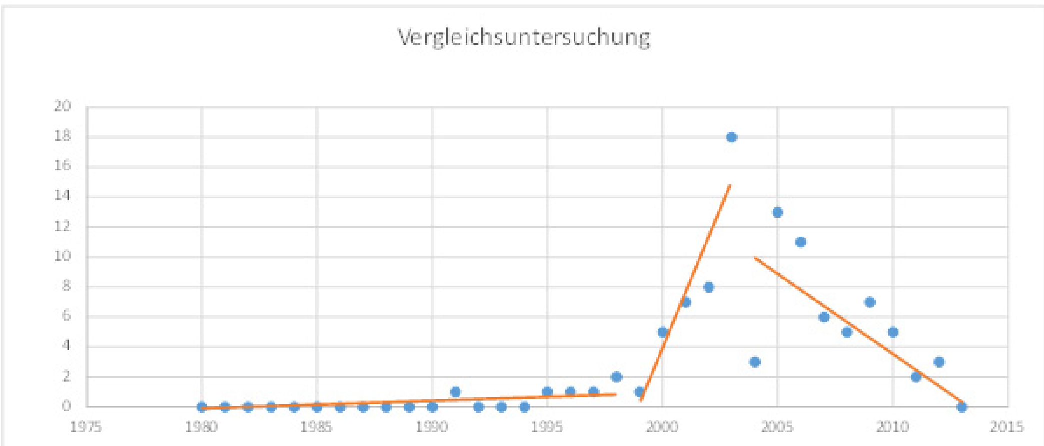


Figure 1. Linear regression segmentation of term "Vergleichsuntersuchung"

possible cutting positions under a certain number of cuts. However the more cuts needed, the higher a punishment factor will be to limit the number of segments.

The above figures illustrates Linear Regression Segmentation of the term “Vergleichsuntersuchung” (comparative study). The straight lines depict the 3 segments of development: 1980-1998, nearly no development; 1999-2003, an abrupt growth; 2004-2013, a steady fall.

Pair of Term Analysis

The behaviour of index term pairs is the focus of our work. We developed several measures to find interesting characteristics from the frequency. As a starting point, we used standard statistical measures such as co-occurrence count and covariance. Then we normalize the co-occurrence count to get co-occurrence ratio. These measures give us many term pairs that develop and change together throughout the years. Then, in order to mine more from the data, we defined several new measures: Deviation from Random, Deviation from Intersection and Co-occurrence Increment Index. They are detailed in the following subsections.

Co-occurrence Count/Ratio

The most direct relationship between two terms is their co-occurrence in a same document. A significant amount of co-occurrence of two arbitrary words can simply show they are more often used together, e.g. "vice" and "versa". But we take only index terms into the processing pipeline, so this meaningless term pairs can be avoided to a great extent. The co-occurrence of index terms can indicate two possibilities: 1. the combination of the two terms forms a topic; 2. the terms are semantically related, like "Schule" (school) and "Schüler" (school student). The first case is obviously useful in trend detection and monitoring. The second case is nevertheless interesting, as it provides a time-based semantic relatedness measure. In comparison to the traditional static semantic relatedness measures, it is based on a time window, in which the relatedness is significant.

The publications are not evenly distributed over the years. So, for each year, we divide the absolute count of the publications containing a specific term by the total count of publications in that year. We call this the co-occurrence ratio of term t_1 and term t_2, denoted by$COR(t_1, t_2, y_1)$. It is defined by the following equation. $\mathrm{Pub}(t_1 \in T, \mathrm{y} = y_1)$ is the set of publication which has t_1 as its index term and is published in year y_1.

$$COR(t_1, t_2, y_1) = \frac{|\mathrm{Pub}(t_1 \in T, \mathrm{y} = y_1) \cap \mathrm{Pub}(t_2 \in T, \mathrm{y} = y_1)|}{|\mathrm{Pub}(t_1 \in T, \mathrm{y} = y_1) \cup \mathrm{Pub}(t_2 \in T, \mathrm{y} = y_1)|}$$

Covariance

Covariance is a popular measure to describe to what extend two random variables tend to change in similar patterns. We take the standard definition of covariance as following and directly apply it to the frequency curves of two terms.

$$cov(X, Y) = \sum_{i=1980}^{2013} \frac{(x_i - \overline{x})(y_i - \overline{y})}{33}$$

Deviation from Random

Our general goal is to find interesting behaviours of the term pairs, co-development is just one kind of that. With the intention not to limit the unknown types of interesting behaviours, we developed this measure to find all types of behaviours that are different from the random distribution. We define the random distribution of co-occurrence of two terms as the null model, which is of least interest. Then we look for those pairs, whose co-occurrence distribution is farthest from the null model. These pairs should possess some strong properties that might be interesting.

The random probability of two terms to co-occur is simply the product of the probabilities of the two terms to occur individually.

$$DFM(X,Y) = \sum_{i=1980}^{2013} \frac{|\mathrm{Pub}(t_1 \in T, \mathrm{y = i})| \cdot |\mathrm{Pub}(t_2 \in T, \mathrm{y = i})|}{|\mathrm{Pub}(t_1 \in T, \mathrm{y = i}) \cap \mathrm{Pub}(t_2 \in T, \mathrm{y = i})| \cdot |\mathrm{Pub}(\mathrm{y = i})|}$$

Here, we use the quotient instead of subtraction to calculate the deviation, as the experiments with quotient show better results.

Deviation from Lower Envelop

Inspired by the results of Deviation from Random, we developed another measure, which we call Deviation from Lower Envelop. It is designed to detect one specific type of term relationship, namely the inter-term dependency. Lower envelop is the curve formed by the lower value of two terms at each time point. As its name suggests, this measure depicts how much the co-occurrence value deviates from the lower value of the two terms at each time point. In other words, for each year, it calculates the difference from the actual co-occurrence value to highest possible co-occurrence value.

$$DFI(X,Y) = \sum_{i=1980}^{2013} \frac{\min(|\mathrm{Pub}(t_1 \in T, \mathrm{y = i})|, |\mathrm{Pub}(t_2 \in T, \mathrm{y = i})|) - |\mathrm{Pub}(t_1 \in T, \mathrm{y = i}) \cap \mathrm{Pub}(t_2 \in T, \mathrm{y = i})|}{33 \cdot |\mathrm{Pub}(\mathrm{y = i})|}$$

By using this measure, we look for the term pairs with smallest DFI value. In these pairs, one term is strongly dependent on the other.

Co-occurrence Increment Index

The trend detection is of high interest. . By its nature, a trendy topic is a topic with increasing popularity. In the preceding measures, the focus lies in co-development, where both increasing and decreasing are considered. Therefore we developed the Co-occurrence Increment Index (CII) is developed to find the term pairs, whose co-occurrence almost monotonically increase. Numerically, this index is the percentage of values in the frequency series that do not violate the monotonic increment of the co-occurrence ratio. The co-occurrence ratio (CR) is co-occurrence at a certain year divided by the lower envelop of that year.

$$CII(t_1, t_2) = \frac{|\{y \in [1980{,}2012] \mid CR(t_1, t_2, y\} < CR(t_1, t_2, y + 1)\}|}{32}$$

Index Terms with Category

In the analysis above, the terms are used without additional information. Our intuitive is terms assigned with categories could provide more insight from the data. In application, it is also easier for users to target the queries to their interested information.

For Pedocs dataset, we manually selected 300 index terms and collaborated with domain experts from DIPF[1] to assign categories (Field, Topic, Method, Geography, Type of Paper, Utility, and Chronicle) to them. With the category, we can look for the term pairs of our interest. For example, we can focus on the method change of topics, by specifying the categories of a term pair into Topic and Method. To investigate the geographical difference of research, we can search the term pair with Field and Geography categories.

---

[1] DIPF: Deutsches Institut für Internationale Pädagogische Forschung (German Institute for Educational Research), Frankfurt am Main, Germany

Empirical Results

Our experimental results show that the fundamental measures, co-occurrence ratio and covariance already produce good results. On the other hand, our novel measures perform also very well at finding interesting characteristics of index term pairs. In this section, we present some representative findings by using different measures.

Co-Occurrence Ratio

Since co-occurrence information is essential to detect any meaningful relationship between index terms, the result from co-occurrence ratio, which derives directly from co-occurrence, is effective at finding related index terms. Figure 2[2] shows "Mathematik" (mathematics) is particularly involved in the research of "Forschendes Lernen" (discovery learning), which may suggest mathematics becomes the main application area of discovery learning.

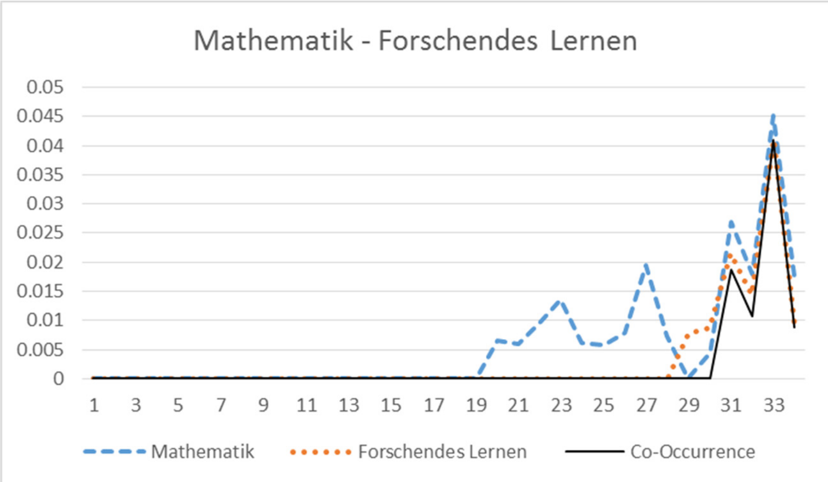


Figure 2. A result of measure Co-Occurrence Ratio

Deviation from Random

Our goal here is to discover the interesting index term pairs, based on their frequency development. Our intuition is that the co-occurrence of index term pairs that follow a random development pattern are not interesting. On the other hand, the pairs whose co-occurrence is far from random are potentially interesting.

In our experiments, we pair the manually selected terms and sort them according to the "Deviation from Random" score. However, pairs with highest score (most deviated from random) usually have a very flat co-occurrence curve, like shown in Figure 3.

Since we define random as the product of frequency of the two terms, it is obvious that a horizontal line deviates extremely from random. They only possible case of higher deviation could only be a curve which develops opposite to the random. Apparently, this is not likely in normal situations. Therefore, the curves most deviated from random are the flat ones.

[2] In this paper, this is the main type of illustration. The horizontal axis is time axis, starting from actual year or year number. The vertical axis is the normalized frequency value, normalized by the total number of publication of a year. There are three curves in the diagram, the dashed curve and dotted curve are the frequency curves of the two index terms, while the curve with solid line is for co-occurrence.

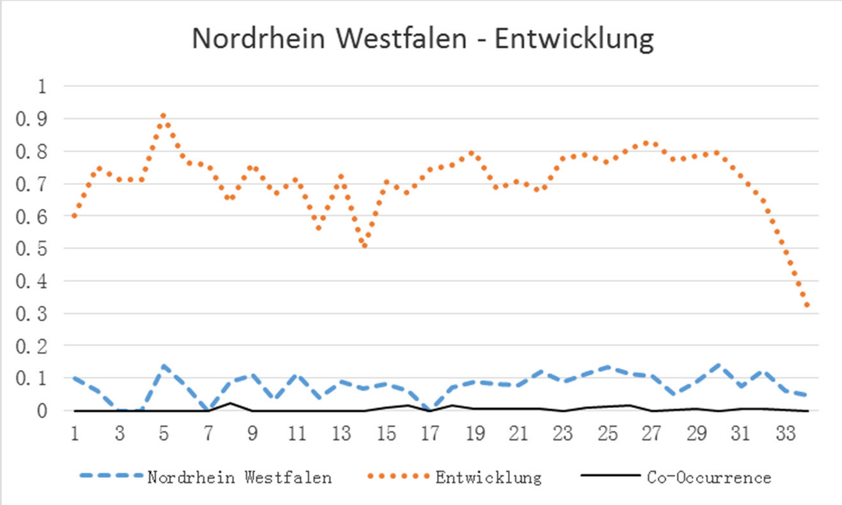


Figure 3. A result from Deviation from Random

This measure can be used for "unrelatedness", as unrelated term pairs have few co-occurrence even when both terms have high frequency.

Deviation from Lower Envelop

This measure calculates the deviation of the co-occurrence from the lower envelop of the two index terms. The high score pairs of this measure have a clear and unique pattern, as shown below:

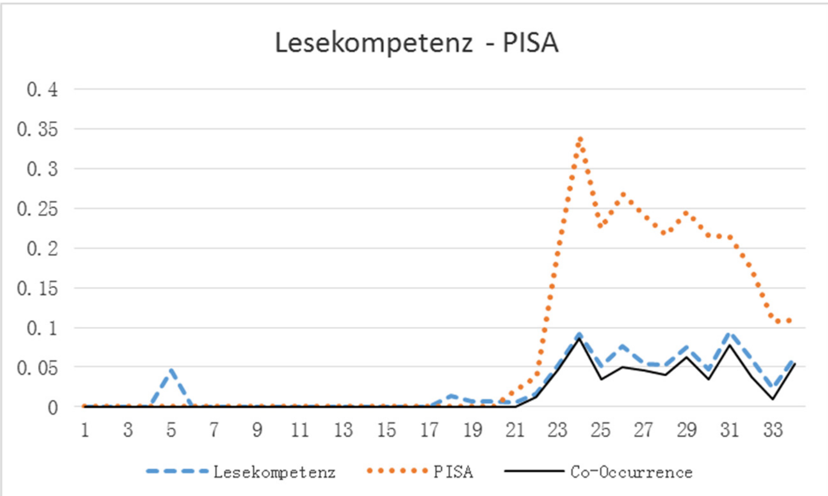


Figure 4. A result from measure Deviation from Lower Envelop

In this example, the co-occurrence curve follows the intersection curve very closely. This depicts a special relationship between the two terms that when one term (Lesekompetenz in this case) is observed in one document, the other term (PISA[3]) is very likely to be observed there too. But the opposite is not true, that when PISA appears in a document, Lesekompetenz is not necessarily likely to appear. This pattern reveals the dependency of one term to another.

Deviation from Lower Envelop with Classified Index Terms

Our experiment results demonstrate temporal analysis on index terms with category information generate more interesting index terms than the one without category information. The examples

---

[3] Programme for International Student Assessment

in Figure 5 are obtained under the condition that the category label of the two index terms are not the same.

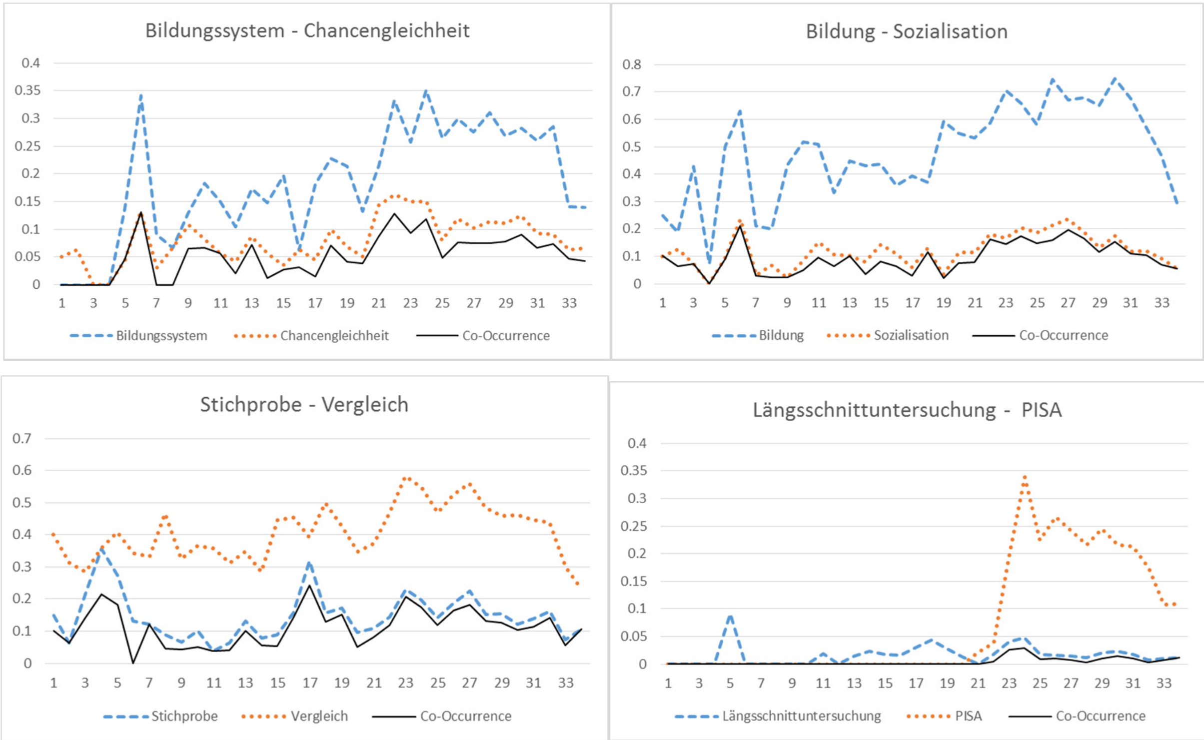


Figure 5. Results from Deviation from Lower Envelop with classified index terms

The upper left example shows “Chancengleichheit” (equal opportunity) is always a concern of Bildungssystem (education system). The upper right diagram demonstrate the dependency of “Sozialisation” (socialization) on “Bildung” (Eductaion). The third diagram on the lower left confirms that “Stichprobe” (sampling) as a method is always used for “Vergleich” (comparison).

The lower right diagram illustrates how one important event in education is reflected by the change in publications. Around year 20 (2001), PISA is introduced. It immediately became the main stream of “Längsschnittuntersuchung” (longitudinal study). In the diagram, we see that the co-occurrence curve follows the lower envelop (the dashed curve for “Längsschnittuntersuchung”) very well.

Acknowledgement

This work has been supported by the Information Center for Education of the German Institute for Educational Research (DIPF) under the Knowledge Discovery in Scientific Literature (KDSL) program.

Conclusion and Outlook

In this paper, we presented our work on temporal analysis of index terms in scientific literature databases. We assigned category information to index terms and obtained better results with it. Our methodology finds interesting pairs of index terms that could be used for domain monitoring and trend detection.

Our next step is to analyse pairs of topics, which are represented by lists of index terms. This would potentially reveal more general temporal events in the database.

References

Allan, J., Carbonell, J., & Doddington, G. (1998). Topic detection and tracking pilot study final report. Retrieved from http://repository.cmu.edu/compsci/341/?utm_source=repository.cmu.edu%2Fcompsci%2F341&utm_medium=PDF&utm_campaign=PDFCoverPages.

Aref, W. G., Elfeky, M. G., & Elmagarmid, A. K. (2004). Incremental, online, and merge mining of partial periodic patterns in time-series databases. *IEEE Transactions on Knowledge and Data Engineering*, *16*, 332–342. doi:10.1109/TKDE.2003.1262186.

Blank, G. D., & Pottenger, W. M. (2001). CIMEL : Constructive , collaborative Inquiry-based Multimedia E-Learning *.

Chan, P. K., & Mahoney, M. V. (2005). Modeling Multiple Time Series for Anomaly Detection. *Fifth IEEE International Conference on Data Mining (ICDM'05)*, 90–97. doi:10.1109/ICDM.2005.101.

Chen, C. (2006). CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature. *Journal of the American Society for Information Science and Technology*, *57*, 359–377. doi:10.1002/asi.20317.

DIPF. FachportalPaedagogik.de. Retrieved from http://www.pedocs.de/.

Havre, S., Hetzler, B., & Nowell, L. (2000). ThemeRiver: visualizing theme changes over time. *IEEE Symposium on Information Visualization 2000. INFOVIS 2000. Proceedings*, 115–123. doi:10.1109/INFVIS.2000.885098.

Keogh, E. (1997). Fast similarity search in the presence of longitudinal scaling in time series databases. *Proceedings Ninth IEEE International Conference on Tools with Artificial Intelligence*. doi:10.1109/TAI.1997.632306.

Kontostathis, A., Galitsky, L., & Pottenger, W. (2004). A survey of emerging trend detection in textual data mining. *Survey of Text …*, 1–44. Retrieved from http://link.springer.com/chapter/10.1007/978-1-4757-4305-0_9.

Lent, B., Agrawal, R., & Srikant, R. (1997). Discovering Trends in Text Databases. *KDD*. Retrieved from http://www.aaai.org/Papers/KDD/1997/KDD97-046.pdf.

Motoyoshi, M., Miura, T., & Watanabe, K. (2002). Mining temporal classes from time series data. *Proceedings of the Eleventh International Conference on Information and Knowledge Management - CIKM '02*, 493. doi:10.1145/584869.584873.

Porter, L., & Detampel, J. (1995). Technology Opportunities Analysis, *10010*, 237–255.

Pottenger, W., Kim, Y., & Meling, D. (2001). HDDI™: Hierarchical distributed dynamic indexing. *Data Mining for Scientific and …*, 1–16. Retrieved from http://link.springer.com/chapter/10.1007/978-1-4615-1733-7_18.

Swan, R., & Jensen, D. (2000). TimeMines: Constructing Timelines with Statistical Models of Word Usage. In *KDD2000 Workshop on Text Mining* (pp. 73–80). Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.17.6612&rep=rep1&type=pdf.

Wei, L., Kumar, N., Lolla, V., & Keogh, E. (2005). Assumption-Free Anomaly Detection in Time Series. *SSDBM*, 1–4. Retrieved from http://www.cs.ucr.edu/~wli/publications/WeiL_AnomalyDetection.doc.

Wimmers, E. L., Psaila, G., Agrawal, R., & Zait, M. (1995). Querying Shapes of Histories. *Proceedings of the 21th International Conference on Very Large Data Bases*, 502–514.

Yang, Z. Y. Z., & Zhao, G. Z. G. (1998). Application of symbolic techniques in detecting determinism in time series [and EMG signal]. *Proceedings of the 20th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. Vol.20 Biomedical Engineering Towards the Year 2000 and Beyond (Cat. No.98CH36286)*, *5*. doi:10.1109/IEMBS.1998.745176.

Posters

List of Accepted Posters

Abdalsamad Keramatfar and Hamzehali Nourmohammadi (Iran)
Area ranking of world's countries according to their rich files in the web

Aliaksei Aldoshyn (Belarus)
Analysis of the Belarusian Universities Positions in the Webometrics Rank

Amir Reza Asnafi (Iran)
A survey on scientific outputs of Humanity faculty members in ShahidBeheshti University (Tehran) during 2006-2010

Amir Reza Asnafi and Maryam Pakdaman Naeini (Iran)
Visibility and co-links analysis of the Iranian scientific associations' websites in the field of Technical and Engineering Sciences

Behjat Taheri and Abdolreza Noroozi Chakoli (Iran)
Comparative analysis of university - industry relation in Iran and Turkey scientometrics study

Behjat Taheri, Hasan Ashrafi-rizi, Leila Shahrzadi and Nayere Sadat Soleimanzade Najafi (Iran)
The Study of outputs and collaboration of Iranian Researchers in disaster in Scopus citation database from 1999 to 2013

Behjat Taheri, Zahra Ghazavi, Faridokht Salahshoori and Hasan Ashrafi-rizi (Iran)
The Scientific Collaboration of the Researchers in Iranian Journal of Medical Education during 2008-2012: A Scientometrics Study

Eunsoo Sohn, Oh-jin Kwon and Jaeyoung Yoo (Korea)
Analysis of the R&D Trend for the Treatment of Alzheimer's Disease through Knowledge Map

Fatemeh Helaliyan Motlagh and Mohammad Hasanzadeh (Iran)
Studying the status of knowledge management

Flor Trillo (Spain)
How does lack of standardization in authorship affect the measurement of science output? Mexican picture

Grant Lewison (UK)
Papers cited by cancer clinical guidelines

Hamideh Asadi (Iran)
Validating scientific social network for Iranian scholars evaluation

Hamzehali Nourmohammadi and Abdalsamad Keramatfar (Iran)
The relation between Wikipedia's size and scientific publication in some countries

Hildrun and Theo Kretschmer (Germany)
Wholeness and Complementary Tendencies in Gender Collaboration

Hongfang Shao, Qi Yu and Zhiguang Duan (China)
Detecting the milestones of epigenetics development from 2002 to 2013 a Scientometrics perspective

Johannes Stegmann (Germany)
Research at UNIS - The University Centre in Svalbard. A bibliometric study

Kyung-Ran Noh, Oh-Jin Kwon and Jae-Young Yoo (Korea)
Scientometric analysis of research activity in EMI shielding fabrics

Madhu Bala, M. P. Singh (India)
A Scientometric Study of Journal of Bio-Chemistry and Bio-Physics (Ijbb)

Maya Verma and Preeti Rani Mishra (India)
Application of Bradford's Law of Scattering to the Literature of History. A Study of Doctoral Theses Citations Submitted to the Chhattisgarh State (India)

Natallia Dudko (Belarus)
Peculiarities of Development of Internet Platforms for Scientific and Technical Production Promotion

Niloofar Hodhodinezhad, Sima Shafian and Hasan Ashrafi Rizi (Iran)
The Study of Scientific Products and Scientific Mapping of Researchers in Disaster during 2002-2012 in the Web of Science (WOS)

Pitambar Gautam (Japan)
Scientific publications on Nepal (up to 2010) in the Web of Science database. A bibliometric visualization

Rajesh Kumar Bhardwaj (India)
Scientometric Dimensions. An Analysis on International Business Literature

Rasool Nouri, Rezvan Ojaghi, Mahmood Keyvan-ara, Urfa Hovsepian and Anasik Lalazarian (Iran)
The researchers' experiences about their ISI publications. A qualitative study

S. Josephine Backiam (India)
Load Balancing in Public Cloud Using Prioritized Round Robin Algorithm

S. S. Pawar (India)
Digital libraries in 21st century and its changing role for Knowledge Sharing /Collaboration. A Study

Sandhya Diwakar, K K Singh and Divya Srivastava (India)
An analysis of research papers (2002-2012) and future perspectives of Nursing in India.

Satish Kumar and Mahender Pratap Singh (India)
A Bibliometric Assessment of International Journal of Special Education

Sedigheh Mohammad Esmaeil and Nader Saboktakin (Iran)
Computer Software Development Challenges In Iran, With Respect To Software Metrics

Sedigheh Mohammad Esmaeil and Zahra Emami (Iran)
Information Retrieval in Endocrine Diseases (Clinical Research) of the World Wide Web

Seyedeh Sara Monir (Germany)
Global trend in h-index research (2005-2014). A bibliometric analysis

Somayeh Ahmadi and Mohammad Hassanzadeh (Iran)
Science Map of Library and Information Science

Vineet Mehan and S. S. Pawar (India)
The seed of Internet

Zahid Ashraf Wani, Tawfeeq Nazir and Muneer Hussain (India)
Visibility of Indian Physics Research. A Citation Study

Zheng Ma and Karsten Weihe (Germany)
Temporal Analysis on Pairs of Classified Index Terms of Literature Databases