



Scientometrics Study on Web: Tools and Techniques

S.Santhanakarthikeyan¹, P. Padma², M.Veeramani³, D.Ravikrishnan⁴

¹Information Officer, Anna Centenary Library, Chennai – 600 085. India.

E-mail: karthikeyan1978@gmail.com

²School of Information and Communication Sciences, Madurai Kamaraj University, Palkalai Nagar,

Madurai – 625 021. E-mail: upswins@gmail.com

³American University of the Middle East, Kuwait,

Email: Veeramani_m@yahoo.com

⁴Librarian and Information Assistant, Anna Centenary Library, Chennai – 600 085.

E-mail: ravibhc@gmail.com

ABSTRACT

This paper aims to discuss the origin and development of scientometric study by comparing and contrasting currently available e-resources with pre internet era. We aspire to illustrate the significance and consequences of e-resources on scientometric study. We shall also briefly touch upon the prospects of fast emerging wireless technology and the necessity of existing methods of study to be constantly transformed to ensure the compatibility of the existing sources with the anticipated technological advancements.

Keywords: Scientometric, Cybermetrics, Bibliometrics

Received 28.02.2013 Accepted 23.03.2013

© 2013 Society of Education, India

INTRODUCTION

Research is vital role for the development of a nation. The research output of a nation is a benchmark to measure its socio-economic and educational status. Many ways to measure the quantity and quality of the research output of the country and even the contributions of an individual. The measurement in the field of library and information science is known as 'bibliometric' analysis. It is interesting to note that during the last few years, bibliometric analysis has been increasingly used to evaluate the research performance of the scientists and the growth of various disciplines of science. The analysis has also been used to evaluate the research output of many leading scientists of world reputations. Before this, it is important to understand the meaning of bibliometrics.

Definitions

Pritchard coined the term 'bibliometrics' and defined it as "The application of mathematical and statistical methods to books and other media of communication". Fairthorne defined bibliometric as: "The quantitative treatment of the properties of recorded discourse and behaviour pertaining to it".

Sengupta viewed scientometric as "Organisation, classification and quantitative evaluation of publication patterns of all macro and micro communications along with their authorship by mathematical and statistical calculus."

Broadus viewed scientometric as "Scientometric is the quantitative study of physical published units or of bibliographic units or of surrogates of either."

Objectives

- ❖ To discuss the origin and development of scientometric study.
- ❖ To illustrate the significance and consequences of e-resources on scientometric study.
- ❖ To compare the existing methods and recent techniques in scientometric study.
- ❖ To explain the various online scientometric databases and its usages.
- ❖ To discuss Sources of Author citation Databases and Databases that provide author citation counts.

Brief historical background of Scientometrics

The origins of bibliometric research can be traced back to the beginning of the 19th century within areas such as law. Shapiro (1999) (2) indicates that many aspects of bibliometrics were “practiced in the legal field long before being introduced into scientific literature”. Early research in the 1880s was reported by Delmas (1992), who describes documentation in France, but initial studies on qualitative and quantitative analysis of science seem to originate within psychological fields (Godin 2006) (3). Godin cites the work of Buchner in describing the notion of “scientific” psychology as “factual, inductive, measurable and experimental” and in 1920 Boring presented research on subject and geographical analysis of psychologists.

Table 1. : Time line of Scientometrics

Early 19th Century	Origins of bibliometric research in areas such as law and psychology	1926-48	Lotka's Law, Zif's Law and Bradford's Law developed.
In 1955	Eugene Garfield first describes the Impact Factor	In 1961	Publication of the Science Citation Index
In 1978	Launch of first dedicated journal, Scientometrics.	1960s-70s	Growth of databases make widespread citation analysis a real possibility

LAWS OF SCIENTOMETRIC STUDY

Probably the earliest, most definable research within the scientometric field was the work that gave rise to the laws of bibliometrics. The first, which came to be known as Lotka’s Law, after Alfred Lotka, can be traced back to 1926 and suggested that within a defined area over a specific period a low number of authors accounted for a large percentage of publications in the area. This was followed in 1935 by the work of George Kingsley Zipf, which describes the frequency of words in a text and became known as Zipf’s Law. Zipf’s research was refined into two main laws looking at high and low frequency words within a text. In 1948 Samuel Clement Bradford’s analysis indicated that within a given area over a specific time a few journals publish a high percent of articles within the area and there are many journals that publish only a few articles each: Bradford’s Law. These laws continue to be studied and form the basis of the development of the modern-day scientometric literature.

In 1944, Lehman described the relationship between quantity and quality within scientific writing and this was followed in 1952 by Dennis, who analysed the effect of scientists’ age on these two elements. Again these types of analyses continue to be described in the current literature, and began to direct thinking towards averaged metrics within bibliometrics.

BIBLIOMETRICS

Bibliometrics is a set of methods to quantitatively analyze scientific and technological literature. Citation analysis and content analysis are commonly used bibliometric methods. While bibliometric methods are most often used in the field of library and information science, bibliometrics have wide applications in other areas. In fact, many research fields use bibliometric methods to explore the impact of their field, the impact of a set of researchers, or the impact of a particular paper. Bibliometrics are now used in quantitative research assessment exercises of academic output which is starting to threaten practice based research. The UK government is considering using bibliometrics as a possible auxiliary tool in its Research Excellence Framework, a process which will assess the quality of the research output of UK universities and on the basis of the assessment results, allocate research funding.

SCIENTOMETRICS

Scientometrics is the science of measuring and analysing science. In practice, scientometrics is often done using bibliometrics which is a measurement of the impact of scientific publications. Modern scientometrics is mostly based on the work of Derek J. de Solla Price and Eugene Garfield. The latter founded the Institute for Scientific Information which is heavily used for scientometric analysis. Methods of research include qualitative, quantitative and computational approaches. One significant finding in the field is a principle of cost escalation to the effect that achieving further findings at a

given level of importance grow exponentially more costly in the expenditure of effort and resources. However, new algorithmic methods in search, machine learning and data mining are showing that is not the case for many information retrieval and extraction based problems. Related fields are the history of science and technology, philosophy of science and sociology of scientific knowledge.

CYBERMETRICS

Cybermetrics is both an Electronic-only Journal and a Virtual Forum (The Journal) devoted to the study of the quantitative analysis of scholarly and scientific communications in the Internet. It is open to world-wide researchers to publish and discuss their findings. Internet offers them new and increased capabilities to distribute their results to a greater audience. Cybermetrics also maintains a series of directories of electronic resources (The Source), including secondary archives of interesting web papers in pdf format. The aim is to provide a reference tool to those researchers involved in the quantitative description and analysis of the Internet as a scholarly communication tool. It is also intended to add original data as a source or reference for larger studies, specially including figures about the distribution and evolution of R&D contents in the World Wide Web.

Cybermetrics focus is on methodologies and results of webometric, scientometric, bibliometric or informetric research with emphasis placed on aspects related to Internet: Application of scientometric methodology to the analysis of scientific communication in the Web and the Usenet, including the development of new Internet R&D indicators. Analyses of hypertext linking phenomena, the informetric laws and distributions, and any mathematical model derived. Impact of the Internet on scientific co-operation and other aspects related to science organization, information flows and interdisciplinary connections. Evaluation of electronic scientific journals and peer-review processes in the World Wide Web. In particular, Cybermetrics is open to those contributions with special requirements, where the new graphical items (multimedia and/or dynamic objects) required by scientometric articles to represent the structure of scientific knowledge production are not suitable for printed publication.

SCIENTOMETRICS ONLINE DATABASE

Special bibliographic database sources are Web of Science, SciVerse Scopus, Compendex, PubMed, etc. Few of the databases are discussed below. The data can be retrieved from these databases for scientometric study in different format. Example .csv, Refworks, Endnote, Tag format, etc. Bibliographical databases such as Web of Science called Science Citation Index, (SCI), Social Science Citation Index (SSCI) and Arts & Humanities Citation Index (A&HCI) maintained by the Institute for Scientific Information (ISI) in Philadelphia, USA. Web of Science covers over 10,000 of the impact journals worldwide, including Open Access journal and over 110,000 conference proceedings and also the retrospective coverage in the sciences, social sciences, arts and humanities available to 1900 (Thomson Reuters. Retrieved 20.3.2011 from http://thomsonreuters.com/products_services/science/science_products/a-z/web_of_science/). Scopus is an international database. It's easy, quick and comprehensive to find the information scientists need. Contains 41 million records, 70% with abstracts, nearly 18,000 titles from 5,000 publishers worldwide, includes over 3 million conference papers, offers sophisticated tools to track, analyze and visualize research (Elsevier BV. Retrieved on 31.3.2011 from <http://www.info.sciverse.com/scopus/about>).

Table 2 : Citation Databases

Sl. No.	Name of Online Database	Specialization	URL	Authority
1	Web of Science	Science, Technology, Social Sciences, Arts & Humanities	http://thomsonreuters.com/products_services/science_products/a-z/web_of_science/	Thomson Reuters
2	Sci Verse Scopus	Science, Technology, Medical, Engineering, Arts & Humanities	http://www.scopus.com/home.url	Science Direct
3	Google scholar	Medical, Scientific, Technical, Business, Social Sciences, Arts & Humanities	http://scholar.google.co.in/	Google

Source : <http://library.amnh.org/research-tools/citation-full-text-databases>

Table 3. : Other Bibliographical and Citation Databases

Sl. No.	Name of Online Database	Specialization	URL	Authority
1	PubMed	Biomedical and Health	http://www.ncbi.nlm.nih.gov/pubmed	US National Library of Medicine National Institutes of Health
2	Compendex	Computer science and Engineering	http://ei.org/compendex	Engineering Village
3	Agricola	Agricultural literature	http://agricola.nal.usda.gov/	National Agricultural Library (USDA)
4	BCIN	conservation and restoration of cultural materials	http://www.bcin.ca/English/home_english.html	BCIN
5	EBSCO Greenfile	All Disciplines	http://web.ebscohost.com/ehost/search/basic?sid=eaed8d77-34bb-4740-a890-14a4638cc54e%40sessionmgr11&vid=1&hid=7	EBSCO Industries
6	TOXNET	Toxicology Literature	http://toxnet.nlm.nih.gov/cgi-bin/sis/htmlgen?TOXLINE	TOXNET

Source : <http://library.amnh.org/research-tools/citation-full-text-databases>

Table 4. : Sources of Author citation Databases

Sl. No.	Name of Online source	Specialization	URL	Authority
1	CiteSeerX	computers in information science	http://citeseerx.ist.psu.edu/	National science foundation
2	ESI (Essential Science Indicators)	Statistics and Science	http://simsrad.net.ocs.mq.edu.au/login?url=http://admin-router.isiknowledge.com/?DestApp=ESI	Thomson Reuters
3	Faculty of 1000	Biology and Medicine	http://f1000.com/prime	F1000
4	Publish or Perish	All Disciplines	http://www.harzing.com/pop.htm#whatfor	Publish or Perish, Inc.

Table 5. : Databases that provide author citation counts

Sl. No.	Name of Online source	Specialization	URL	Authority
1	ACM Digital Library	Computing Literature	http://simsrad.net.ocs.mq.edu.au/login?url=http://portal.acm.org/dl.cfm?coll=portal&dl=ACM&CFID=1436618&CFTOKEN=5260756	ACMDL
2	EBSCOhost databases	All Disciplines	http://simsrad.net.ocs.mq.edu.au/login?url=http://search.epnet.com/	EBSCO
3	JSTOR	All Disciplines	http://simsrad.net.ocs.mq.edu.au/login?url=http://www.jstor.org/search/AdvancedSearch	ITHAKA
4	MathSciNet	Mathematics	http://simsrad.net.ocs.mq.edu.au/login?url=http://www.ams.org/mathscinet/	American Mathematical Society
5	Ovid	Medicine, Health, Humanities and Behavioural Sciences	http://simsrad.net.ocs.mq.edu.au/login?url=http://ovidsp.ovid.com/autologin.html	Ovid Technologies
6	PubMed	Medicine	http://www.ncbi.nlm.nih.gov/sites/entrez?db=pubmed	NCBI
7	SciFinderScholar	Science	http://www.library.mq.edu.au/research/databases/scifinderweb.html	American Chemical Society
8	SpringerLink	All Disciplines	http://simsrad.net.ocs.mq.edu.au/login?url=http://www.springerlink.com/app/home/main.asp?wap=215lxjuumg7qd7nwdrwq	Springer

Source : <http://libguides.mq.edu.au/content.php?pid=115026&sid=1151527#Other>
author citation databases

Table 6. : Scientometrics tools

Sl. No.	Name of the Tool	Purpose	URL	Type	Source	Compatibility
1	Authormap	Citation Mapping and Visualization	http://project.cis.drexel.edu/authorlink/	Web-tool	Howard White, et. Al. Drexel University, hdwhite@drexel.ed	N/A, FlashPlayer Required
2	Bibcouple.exe	Visualization of the bibliographic coupling among authors using WoS set	http://users.fmg.uva.nl/lleydesdorff/software/bibcoupl/index.htm	Software Application	Loet Leydesdorff	PC / DOS. Works with Pajek, MS Access and Excel
3	BibJourn.exe	Visualization of the bibliographic coupling in terms of cited journals	http://users.fmg.uva.nl/lleydesdorff/software/bibjourn/index.htm	Software Application	Loet Leydesdorff	PC / DOS. Works with Pajek and MS Access and Excel
4	Citespace	Visualizing patterns and trends in scientific literatur	http://cluster.cis.drexel.edu/%7Ecchen/citespace/	Map	Chaomei Chen,	Images: N/A application: Java required
5	CleanPoP	Tool is designed to clean results systematically. Publish Or Perish to	http://cleanpop.ifris.org/guide.htm	Web-tool	Audrey Baneyx/ IFRIS	better with firefox 3 / use browser that respect W3C
6	Co-auth.exe	Visualization of the coauthorship network using a WoS set	http://users.fmg.uva.nl/lleydesdorff/software/coauth/index.htm	Software Application	Loet Leydesdorff	PC / DOS. Works with Pajek and MS Access and Excel
7	Fulltext	Software for co-word mapping of full text	http://users.fmg.uva.nl/lleydesdorff/software/fulltext/index.htm	Software Application	Loet Leydesdorff	PC / DOS. Works with Pajek and MS Excel
8	HistCite	Bibliographic Analysis and Visualization Software	http://www.histcite.com/index.htm	Software Application	Dr Eugene Garfield, founder of the Institute for Scientific Information and the inventor of the Science Citation Index	PC
9	IntColl.exe	For Visualization of international collaboration	http://users.fmg.uva.nl/lleydesdorff/software/intcoll/index.htm	Software Application	Loet Leydesdorff	PC / DOS. Works with Pajek and MS Access and Excel
10	ISI	For organizing a set downloaded from the Web of-Science into databases for relational database management	http://users.fmg.uva.nl/lleydesdorff/software/isi/index.htm	Software Application	Loet Leydesdorff	PC / DOS. Works with Pajek and MS Access and Excel
11	Patent Pictures	It's patently good news	http://www.researchinformation.info/rijanfeb04patents.html	Software Application	Research Information	N/A
Sl. No.	Name of the Tool	Purpose	URL	Type	Source	Compatibility

12	Publish or Perish	Retrieves and analyzes academic citations from Google Scholar	http://www.harzing.com/pop.htm	Software Application	Google Scholar	N/A
13	RefViz	Data visualization and analysis software from the makers of EndNote, ProCite, and Reference Manager for exploring reference collections based on content	http://www.refviz.com	Software Application	Thomson ResearchSof	Mac and PC. Interface with EndNote, ProCite, Reference Manage
14	TL.exe	Co-word mapping of texts	http://users.fmg.uva.nl/lleydesdorff/software/ti/index.ht	Software Application	Loet Leydesdorff	PC / DOS. Works with Pajek and MS Excel

Source : <http://www.demoscience.org/resources/category/76>

CONCLUSION

The study gives interesting and important information about scientometric study and its tools available in online. Important tools and techniques of scientometric study for areas of scientometric research. It is hoped that this study will be helpful to researchers who want to identify primary sources of scientometric study. Studies of this kind will be helpful for library and information professionals who want to provide suitable online sources for users and researchers.

REFERENCES

1. Abercrombie, Robert K. and Udoeyop, Akaninyene W.,(2012). A Study of Scientometric Methods to Identify Emerging Technologies. New York: U.S. Department of Energy.
2. Abercrombie, Robert K., Udoeyop, Akaninyene W. and Schlicher, Bob G., (2012). A study of scientometric methods to identify emerging technologies via modeling of milestones. *Scientometrics*. 91 (2). : 327-342. New York: Springer-Verlag..
3. Karpagam,R, Gopalakrishnan, S., and Natarajan, M., (2011). Scientific measures and tools for research literature output. *Indian Journal of Science and Technology*. 4 (7): 828-833. Chennai: Indian Society for Education and Environment.
4. Zhao, Dangzhi., Logan, Elisabeth.,(2002). Citation analysis using scientific publications on the Web as data source: A case study in the XML research area. *Scientometrics*, 54 (3): 449-472.New York: Springer-Verlag.
5. Mike Thelwall (2009). Introduction to Webometrics: Quantitative Web Research for the Social Sciences. Morgan & Claypool. ISBN 978-1-59829-993-9.
6. <http://libguides.mq.edu.au/content.php?pid=115026&sid=1151527#Otherauthor> citation databases
7. <http://library.amnh.org/research-tools/citation-full-text-databases>
8. <http://www.demoscience.org/resources/category/76>
9. <http://library.amnh.org/research-tools/citation-full-text-databases>