A Scientometric Portrait of the Journal *Digital Investigation*

J. John Jeyasekar
Librarian,
Forensic Sciences Department,
Mylapore, Chennai – 600 004
(Part-Time Research Scholar,
Manonmaniam Sundaranar University)
mail: trijays@gmail.com

P. Saravanan
Librarian
Lekshmipuram College of Arts & Science,
Kanyakumari District
mail: sara_nps@yahoo.co.in

Abstract

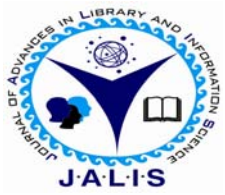
A scientometric study of the journal Digital Investigation for the period 2001 to 2013 was done using data obtained from Google Scholar. Publish or Perish Software was used for the retrieval and analysis of the data and VOS viewer software was used for cluster analysis and visualization. The growth pattern of literature, authorship pattern and citations were studied. The study revealed that the literature growth followed a linear pattern. The degree of collaboration was 0.46. The h-index and g-index were 39 and 56 respectively

Keywords

Scientometrics, Literature Growth, Authorship Pattern, Citation Study, Single Journal Study, Digital Investigation.

Electronic access

The journal is available at www.jalis.in



Journal of Advances in Library and Information Science
ISSN: 2277-2219 Vol. 3. No.2. 2014. p155-162

INTRODUCTION

Forensic science may be defined broadly as the application of science to the analysis and interpretation of physical evidence in criminal and civil litigation. Forensic science refers to the application of principles and methods of specialized scientific and technical knowledge to criminal and civil legal questions and presenting the finding in an unbiased and objective way in courts of law. According to Saferstein (2001) "Forensic science is the application of science to those criminal and civil laws that are enforced by police agencies in a criminal justice system." Thus forensic science is a science related to the police agencies and to the judiciary.

Forensic science is more than a collection of fragments borrowed from the various basic disciplines from astronomy to zoology. Forensic sciences include, but are not limited to pathology, psychiatry, psychology, odontology, toxicology, molecular biology, entomology. Digital forensics also known as digital forensic science is a branch of forensic science encompasses the recovery and investigation of material found in digital devices, often in relation to computer crime. The term digital forensics originally used as a synonym for computer forensics with the development in Information and Computer Technologies cover investigation of all devices capable of storing of digital data. The technical aspects of digital forensics include several sub-branches relating to the type of digital devices involved; computer forensics, network forensics, forensic data analysis, and mobile device forensics. The typical digital forensics involves the seizure, forensic imaging acquisition, and analysis of digital media and the production of a report into collected evidence.

SCIENTOMETRICS

Scientific research is a large enterprise and the substance of it is so complex and specialized that personal knowledge and experience are no longer sufficient to measure the trends and quantify the results. A government office weighing national research need, a funding agency making a funding choice, and a library facing collection related policy decision have to depend on the evaluation of scientific research performance. Peer review was the standard approach to evaluate a research project until partially being replaced by bibliometric tools such as publication counts and citation analyses. Peer review,

i.e., experts reviewing the works of their colleagues should rightly be the basis of research evaluation. However, there are certain flaws like bias and favoritism inherent in this system. Hence bibliometric / scientometric tools are increasingly being used for the purpose.

The term 'bibliometrics' was first coined by Alan Pritchard and it is defined as "application of mathematics and statistical methods to books and other media of communication" (Pritchard, 1969). Roy (1983) has defined bibliometrics as a "study of the process of information use by analyzing the characteristics of documents and their distribution by statistical methods". The origin of the term 'scientometrics' dates back to the year 1969, when the Russian scientists Nalimov and Mulechenko coined the Russian term 'naukometriya' the Russian equivalent of scientometrics (Nalimov and Mulechenko, 1969). However, the advent of scientometrics as a discipline was in 1978, when the journal 'Scientometrics' was founded by Tibor Braun. Scientometrics defines its content as "Scientometrics includes all quantitative aspects of the science of science, communication in science, and science policy." (Wilson, 1999) The focus of Scientometrics is the measurement of science and is therefore concerned with the growth, structure, interrelationship and productivity of scientific disciplines. Tague-Sutcliffe (1992) defines "Scientometrics is the study of the quantitative aspects of science as a discipline or economic activity. It is part of the sociology of science and has application to science policy-making. It involves quantitative studies of scientific activities, including, among others, publication, and so overlaps bibliometrics to some extent." Scientometric study of a single journal helps to create a portrait of the journal. It gives a description that offers insight beneath the superficial. It indicates the quality, maturity and productivity of the journal in the particular field and geographical region. The journal thus studied is regarded as the most important in the given subject area.

THE JOURNAL *DIGITAL INVESTIGATION*

The journal *Digital Investigation* covers cutting edge developments in digital forensic and incident response from around the globe. This widely referenced publication helps investigators remain current on new technologies, useful tools, relevant research, investigative techniques, and methods for handling security breaches. Elsevier, the publisher of this periodical started publishing the journal in the year 2004. Till 2010, ten volumes have been published.

NEED AND SIGNIFICANCE OF THE STUDY

Scientometric studies have increasingly been used over the last few years. These studies are useful to understand the evolution of literature or trends in particular fields or within a geographical area. However, in forensic science, scientometrics have barely been used. Alan Wayne Jones has worked on bibliometric analysis of Forensic science literature. His interesting work is mainly focused on most highly cited articles, most prolific authors and impact factors (Sauvageau, Desnoyers and Godin, 2009). Jeyasekar and Saravanan have published a few bibliometric and scientometric studies in the area of forensic science. Other than these studies there are no much studies in this field. Hence, scientometric studies in forensic science have considerable significance to understand the latest trends. Several single journal studies have been conducted in all the domains of knowledge as discussed elsewhere in this paper. The present study is a scientometric study of a single journal in the field of forensic science.

OBJECTIVES OF THE STUDY

This study is intended to analyse the

- Growth of literature;
- Authorship pattern and the degree of collaboration;
- Citations; and
- Cluster analysis and visualization of the journal *Digital Investigation* for the 10 year period of 2004 to 2013 covering the first 10 volumes of the journal.

REVIEW OF LITERATURE

Several studies have been conducted to study a particular journal with regard to its authorship pattern, publication pattern, authorship collaboration, and so on. According to Anyi, Zainab and Anuar (2009) a total of one hundred and two bibliometric studies with a single scholarly journal have been published between 1969 and 1997 and Eighty two bibliometric studies on single journals have been published between 1998 and 2008 [3]. However, in the field of Forensic Science, single journal studies are uncommon. The two single journal studies in the field of forensic science are as follows.

Jones (1998) studied the citation pattern of the '*Journal of Forensic Sciences*' and states that Government forensic scientists probably care little about the notion of journal impact factor because their promotional prospects and salary increments are not connected to scientific productivity as reflected in their publication track record. Jeyasekar and Saravanan (2013) carried out a bibliometric study of the '*Journal of Forensic Sciences'* and found that there is an increase in publications on digital and multimedia aspects of forensic science and the literature related to application of DNA technology in forensic science is also increasing. The mean degree of authorship collaboration of the Journal of Forensic Sciences is 0.91. The present study is a scientometric analysis of the journal *Digital Investigation* based on the publication data available in Google Scholar using Publish or Perish (PoP) software.

MATERIALS AND METHOD

Data for the present study are retrieved from Google Scholar (GS). An important advantage of GS is that it is freely available to anyone with an internet connection and is generally praised for its speed. Other bibliographic data sources such as SCOPUS and Web of Science (WoS) are available to those academics whose institutions are able and willing to bear the high subscription cost. GS is not designed as a direct competitor to the other citation database. But it is really a huge database and Google is clearly intending to enlarge its coverage. GS can be used for bibliometric evaluation but with great care (Aguillo, 2011). The data obtained are analysed using Harzing's Publish or Perish (PoP) Software programs. PoP is a free software program that retrieves and analyses academic citations. It uses Google Scholar to extract the raw citations and analyses these and presents the statistics such as total number of papers, total number of citations, average number of citations per paper, average number of citations per author, average number of papers per author, average number of citations per year, h-index, g-index, contemporary h-index, the age-weighted citation rate, two variations of individual h-indices and analysis of number of authors per paper. The results can easily be copied to various Windows application.

The present study used 'Digital Investigation' as keyword in the journal query option and the year was limited to each year starting from 2004 to 2013. The above exercise was done on 20th March 2014. The data was downloaded in MS-excel format and analysed.

RESULTS AND DISCUSSION

Growth of Literature

(i) Year-Wise Growth

The year-wise number of papers, their percentage, cumulative growth of papers, cumulative growth percentage, Relative Growth Rate (RGR) and Doubling Time (Dt) are given in Table 1.

Table 1: Growth of literature

Year	No of papers	Per cent	Cumulative growth	Per cent	RGR	Mean RGR	Dt	Mean Dt
2004	44	7.75	44	7.75				
2005	58	10.21	102	17.96	0.84		0.83	
2006	71	12.50	173	30.46	0.53		1.31	
2007	68	11.97	241	42.43	0.33		2.10	
2008	30	5.28	271	47.71	0.12	0.36	5.78	2.00
2009	39	6.87	310	54.58	0.14		4.95	
2010	60	10.56	370	65.14	0.17		4.08	
2011	42	7.39	412	72.54	0.11		6.30	
2012	71	12.50	483	85.04	0.16		4.3	
2013	85	14.96	568	100	0.16	0.15	4.3	4.79
	568	99.99						

The journal has published a total of 568 papers during the period of study. The journal started with 44 papers in its first volume. The number of papers grew steadily to 58 and 71 in the next two years. The number of papers then declined to 68 and 30 in the next couple of years and again grew steadily to 85 in the year 2013. However, it declined to 42 in between

in the year 2011. The year-wise publication pattern is depicted in figure 1.

Figure 1: Year-wise growth of papers

(ii) Cumulative Growth

The cumulative growth pattern has seen a steady increasing linear trend in a straight line rather than the curve pattern of exponential trend propounded by Price (1969).This is illustrated in Figure 2. Relative Growth Rate (RGR) is the measure to study the increase in number of papers of a particular period of time. Doubling Time (Dt) is the period of time required for a quantity to double in size or value (Karpagam, et al. 2011). Dt is inversely related to RGR. The mean RGR of the first 5 years period is 0.36 and the second 5 years period is 0.15. Since Dt and RGR are inversely related it is observed that the mean Dt which is 2.00 in the first 5 years period grew to 4.79 in the second 5 years period.

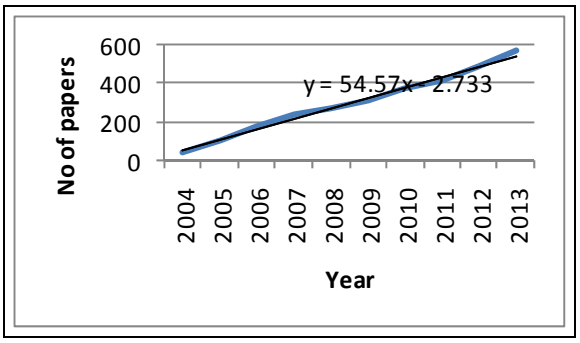


Figure 2: Cumulative growth pattern

Authorship Pattern

The authorship pattern is enumerated in Table 2.

Table 2: Authorship pattern

Author	No of papers	Per cent
Single	304	53.52
Two	108	19.01
Three	93	16.37
Four	37	6.51
Five	16	2.82
Six	9	1.58
Seven	1	0.18
Total	568	99.99

Scientific research is a co-operative enterprise. Hence the research output in the written form of communication shall also have more authors. Studying the authorship pattern helps us to get a fair idea of the amount of co-operation of the researchers. In the present study it is found that 304 papers (54% of the total) are contributed by a sole author and 264 papers (46% of the total) are contributed by multiple authors. The numbers of papers contributed by two authors are 108 (19%); three authors are 93 (16%); four authors are 37 (7%); and five authors are 16 (3%).

Numerous studies to find the degree of collaboration and collaboration co-efficient have been conducted. Research activity in a multidisciplinary subject like forensic science necessarily has to be highly collaborative one. In the case of this journal publication records during the period of study it is found that only 46 per cent of the papers are multi-authored ones. The authorship pattern is illustrated in Figure 3.

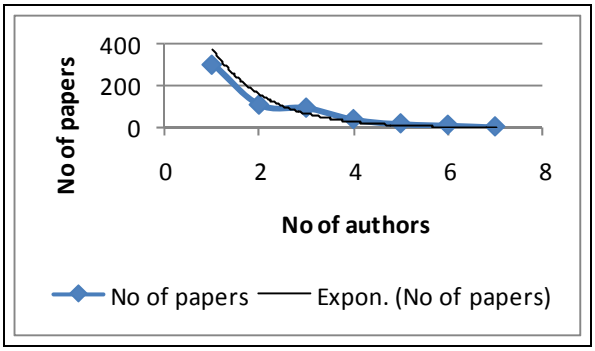


Figure 3: Authorship pattern

Degree of Collaboration

Subramanyam (1983) enunciated the formula C=Nm/(Ns+Nm) , where C is the degree of collaboration, Nm is the number of multi-authored papers and Ns is the number of single authored papers. The degree of collaboration in this journal as per this formula is 0.46. This simply means that 46

percent of the papers have more than one author, which can also be understood from the data available in Table 3. This is in contrast to the degree of collaboration found in the 'Journal of Forensic Sciences', which was as high as 0.91. It is quite natural that the degree of collaboration in forensic sciences to be very high because research in this field requires understanding and application of various sciences. Nevertheless digital forensics being a very narrow field of forensic sciences much collaboration is not required and hence the low degree of collaboration is justified.

Citation Analysis

i) Growth

The journal has received a total of 5748 citations during the period of study. The year-wise break-up of the citations and the cumulative growth of citations are given in Table 3.

Table 3: Growth pattern of citation

Year	Total no of citation	%	Cumula-tive growth	RCR
2004	711	12.37	711	1.24
2005	570	9.92	1281	0.99
2006	1178	20.49	2459	2.05
2007	814	14.16	3273	1.42
2008	679	11.81	3952	1.18
2009	543	9.45	4495	0.94
2010	621	10.80	5116	1.08
2011	307	5.34	5423	0.53
2012	270	4.70	5693	0.47
2013	55	0.96	5748	0.10
Total	5748	100		

It is observed that the citations received does not have definite pattern. It is normal for the citations to be low in the latest years. Hence the low citation rate during the year 2013 is justifiable. However during the year 2006 a sudden spurt in the number of citations is found and number of citations peaked in that year. If Impact Factor (IF) is calculated using a 5 year window instead of the present practice of using a two year window may lead to higher IF and also a more realistic evaluation of journals.

ii) Relative Citation Rate

Relative Citation Rate (RCR) is the ratio between the Observed Citation Rate and Expected Citation Rate. This concept was first used to calculate the RCR of a given geographical area. However, it is applied here to study the citation trend of a particular journal. The RCR is said to be normal when the score is 1 and below average when it is below 1 and above normal when it is above 1. The journal *Digital Investigation* has normal RCR in the year 2005 and above average during the years 2004, 2006, 2007, 2008 and 2010. RCR is below average during the years 2009, 2011, 2012 and 2013. It is also found that RCR is as high as 2.05 in the year 2006.

iii) Papers – Citations Ratio

Sagar, et al (2010) observed that more number of papers in a particular year received more number of citations and this indicates that quality and quantity of research always go hand in hand. However, in the present study no such relationship is observed.

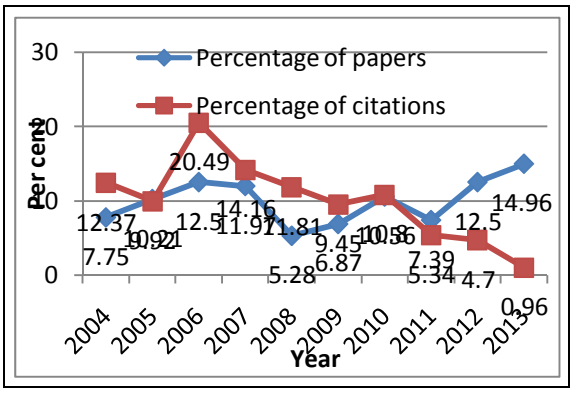


Fig 4: No of papers – citations ratio

iv) h-Index

Hirsch (2005) introduced the h-index to measure both the productivity and impact of the published work of a scientist. The index is based on the set of the scientist's most cited papers and the number of citations that they have received in other people's publications. h-index is defined as the number of papers (h) with citation number higher or equal to h. i.e., "a scientist has index h, if h of his/ her N papers have at least h citations each and the other (N-h) papers have no more than h citations. Hirsch also states that h-index does measure the broad impact of individual's work. Braun, Glanzel and Schubert applied this concept of computing h-index for journals. The index can also be applied to the productivity and impact of a group of scientists, such as a department or university or country. This has evolved into an index to measure the productivity and impact of a particular journal too.

The h-index of *Digital Investigation* obtained from Google Scholar during the period of study is 39. This means that 39 papers have received 39 or more citations during the study period. The year-wise distribution of these 39 h-index papers and the total citations received these papers are given Table 4.

Table 4: Distribution of h-index papers

Year	No of h-index papers	Total citations
2004	5	403
2005	4	288
2006	9	679
2007	6	357
2008	9	470
2009	2	134
2010	3	228
2011	1	49
Total	39	2608

It is observed that the numbers of h-index papers are more during the years 2006 and 2008, where 9 papers in each year received as many as 679 and 470 citations respectively. In 2007, 6 papers come under this category but with only 357 citations whereas in 2004, 5 papers with a total citation of 403 among them come under the category of h-index papers. 2011 has the lowest number of h-index papers and citations with 1 and 49 respectively.

v) g-index

Egghe (2006) introduced the g-index as an improvement of Hirsh's h-index for measuring the overall citation record of a set of papers. The influence of the self citations appears to be more significant for the g-index than the h-index. The g-index of the journal *Digital Investigation* during the period of study is 56.

v) Top Cited Papers

Six papers have received more than 100 citations per paper and these papers are listed in Table 5. Out of these six papers two are papers published during the year 2006 and this explains the citation peaking during the year 2006.

Table 5: Papers with very high h-index

Cites	Authors	Title	Year
154	BD Carrier, J Grand	A hardware-based memory acquisition procedure for digital investigations	2004
141	NL Beebe, JG Clark	A hierarchical, objectives-based framework for the digital investigations process	2005
140	SL Garfinkel	Digital forensics research: The next 10 years	2010
134	A Schuster	Searching for processes and threads in Microsoft Windows memory dumps	2006
126	J Kornblum	Identifying almost identical files using context triggered piecewise hashing	2006
121	SL Garfinkel	Carving contiguous and fragmented files with fast object validation	2007

Cluster Map

Analysis of words used in the title of the papers is done using VOSviewer. The word clusters obtained are listed in Table 6 and the map obtained is given Figure 5.

Table 6: Clusters

Cluster # 1		Cluster # 2		Cluster # 3	
Terms	Frequency	Terms	Frequency	Terms	Frequency
Case study	7	Evidence	23	Forensic	59
Digital investigation	11	Investigation	34	Tool	28
Forensic analysis	15	Need	8		
State	6				

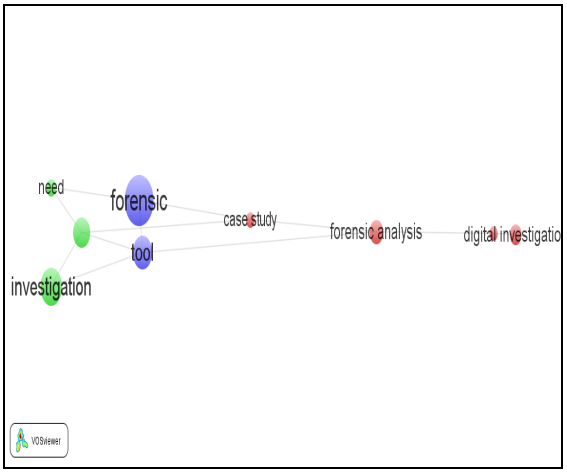


Figure 5: Cluster Map

CONCLUSION

Digital Investigation, an important journal in the field of forensic sciences dealing with the most current area of field has published 568 papers during the period of study. This journal received 5748 citations during that period. The RGR during the first half period is 0.36 which decreased to 0.15 during the second half period. The Dt during the first half period is 2.00 and the second half period is 4.79. The degree of collaboration is 0.46. The number of citations peaked to the maximum of 1178 in the year 2006. The h-index and g-index of the journal during the period of study are 39 and 56 respectively.

REFERENCES

1. Aguillo, I. F. (2011). Is Google Scholar Useful for Bibliometrics? A Webometric Analysis. Scientometrics. doi 10.1007/s11192-011-0582-8

2. Anyi, K. W. U., Zainab, A. N., & Anuar, N. B. (2009). Bibliometric Studies on Single Journals: A Review. Malaysian Journal of Library & Information Science, 14(1), 17-55.

3. Egghe, L. (2006). Theory and Practice of the g-index. Scientometrics, 69(1), 131-152.

4. Harzing, A. W. (2013). Google Scholar - A new data source for citation analysis. Retrieved from http://harzing.com/pop.html

5. Harzing, A. W. (2007). Publish or Perish. Retrieved from http://www.harzing.com/pop.htm

6. Hirsch, J. E. (2005). An Index to Quantify an Individual's Scientific Research Output. Proceedings of the National Academy of Sciences, 102, 16569-16572. Retrieved from http://www.pnas.org/cgi/doi/10.1073/pnas.0507655102

7. Holt, C. (2006). Guide to Information Sources in the Forensic Science. Connecticut: Libraries Unlimited.

8. Jeyasekar, J. J., & Saravanan, P. (2012). Scientometric Analysis of Forensic Science Publications: A study based on SCOPUS database. UGC-SAP National Seminar on Scientometrics and Informetrics (pp. 75-78). Annamalai Nagar, India: Annamalai University.

9. Jeyasekar, J. J., & Saravanan, P. (2013). Journal of Forensic Sciences: A Bibliometric Study for the Period 2006 to 2010. Second National Conference on Scientometrics and Knowledge Management. Dharwad, India.

10. Jones, A. W. (1998). Citation Trends and Practices in the Journal of Forensic Sciences as Documentated by ISI's Journal Citation Report. Journal of Forensic Sciences, 43(2), 439-444.

11. Jones, A. W. (2003). Impact Factors of Forensic Science and Toxicology Journals: What Do the Numbers Really Mean. Forensic Science International, 133(1), 1-8.

12. Jones, A. W. (2005). Which Articles and Which Topics in the Forensic Sciences are Most Highly Cited? Science & Justice, 45(3), 175-182.

13. Jones, A. W. (2007). The Distribution of Forensic Science Journals, Reflections on Authorship Practices, Peer-review and the Role of the Impact Factor. Forensic Science International, 165(2), 115-128.

14. Karpagam, R., Gopalakrishnan, S., Natarajan, M., & Ramesh Babu, B. (2011). Mapping of nanoscience and nanotechnology research in India: a scientometric analysis, 1990-2009. Scientometrics, 89, 501-522.

15. Nalimov, V. V., & Mulechenko, Z. M. (1969). Naukometriya Izuxhenie Razvitiya Nauki kak Informasionnog Protsessa. Moskow: Nauka.

16. Price, de Solla (1963). *Little Science, Big Science*. New York: Columbia University Press.

17. Pritchard, A. (1969). Statistical bibliography or bibliometrics. Journal of Documentation, 25(4), 348-349.

18. Roy, P. M. (1983). Towards a theory of citing in citation analysis studies, In: C.D. Sharma and Kailash Vyas, eds. Developing horizons in Library and Information Science, Vol. 1. Jaipur: Print Well.

19. Saferstein, R. (2001). Criminalistics: An Introduction to Forensic Science (7 ed.). New Jersey: Prentice-Hall.

20. Sagar, A., Kademani, B. S., Garg, R. G., & Kumar, Vijai (2010). Scientometric Mapping of Tsunami Publications: A Citation Based Study. Malaysian Journal of Library and Information Science, 15(1), 23-40.

21. Sauvageau, A., Desnoyers, S., & Godin, A. (2009). Mapping the Literature in Forensic Science: A Bibliomeric Study of North American Journals from 1980 to 2005. The Open Source Forensic Science Journal, 2, 41-46.

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

23. Tague-Sutcliffe, J. (1992). An Introduction to Informetrics. Information Processing and Management, 28(1), 13.

24. VOSviewer retrieved from http:// www.vosviewer.com

25. Wilson, C. S. (1999). Informetrics. Annual Review of Information Science and Technology, 34, 107-247.