OPEN SOURCE SOFTWARE (OSS): A SCIENTOMETRIC STUDY OF GLOBAL PUBLICATIONS DURING 1999-18

Jivesh Bansal
Madhu Bansal

Jivesh Bansal
Panjab University, A.C.Joshi Library, Chandigarh 160014
jivesh@pu.ac.in

and

Madhu Bansal
Panjab University,
Department of Mathematics
Library, Chandigarh 160014
madhu@pu.ac.in
Corresponding Author:
Madhu Bansal

The paper examines 341 global publications on "Open Source Software" research, as seen from Web of Science bibliographical database during 1999-18. The study uses various bibliometric indicators, such as growth rate, citation impact, share of international collaborative papers, subject distribution etc to explore research in this area. The OSS research registered 23.12% growth and registered average citation impact per paper of 7.54 citations. The global share of top 10 most productive countries in OSS research ranged from 2.05% to 32.55%. USA contributed the largest global share (32.55%), followed by India (10.26%), England (7.62%), China (5.87%), Spain (5.28%), Canada (3.81%) etc. Six countries scored relative citation index above the world average of 1.0: Australia (1.81), England (1.55), China and Spain (1.39 each), USA (1.117) and Canada (1.08) during 1999-18. The share of international collaborative publications of top 10 most productive countries in OSS research varied from 2.86% to 87.50%. Information Science Library Science (100%), accounted for the highest publication share among broad subjects, followed by Computer Science Information Systems (30.79%), management (10.26%), communication (3.52%) etc. during 1999-18. The top 15 most productive organisations and authors contributing to OSS research accounted for 19.35% and 12.02% global publication share and 18.9% and 11.66% global citation share, respectively. The top 15 most productive journals accounted for 61% global share of journal publication output during 1999-18.

Keywords: Open Source Software, Global publications, Bibliometrics, Scientometrics

INTRODUCTION

Information technology has revolutionized the services and functioning of libraries. Technological developments, like use of software have expedited the management and access to huge information. Library management software, content management systems and digital library software are various kinds of library software used for library automation and creation of institutional repositories or digital libraries. Library

management software ease the arduous routine of library services with its different modules, like accessioning, circulation, serials control, updated bibliographic records, OPAC etc. Use of content management systems and digital library software help to disseminate scholarly research and manage electronic resources, provide an advanced discovery interface[1]. Library software can be either proprietary or open source. Proprietary software are obtained by purchase and charge huge annual maintenance charges, frequent revisions and up-gradations, lack of support for previous versions and without freedom to modify the software. Exorbitant prices of commercial software and ever shrinking budgets of libraries have put the libraries in a quandary.

Open Source Software have emerged as a panacea for institutions and libraries. Open Source Software are called free software/ libre software, FOSS, FLOSS or OSS. Open Source Software (OSS) licenses the user to run the program, modify and redistribute the copies of modified or original program by the availability of its source code. Libraries can modify the program to suit its needs and improvise its services. Open Source Software is a boon for libraries in underdeveloped or unprivileged regions[2,3]. Libraries are saved from installation and maintenance cost of proprietary software with their better functionality and quality. Open Source Software selected for use should be freely downloadable, standard based and in compatibility with operating systems like Linux, Windows etc. Availability of source code in open domain leads to their sharing and usability much easier[4,5].

Widespread usage of OSS in libraries have led to provision of instantaneous access to information and better services to patrons and helped in bridging the digital divide. Zero cost of acquisition, low maintenance costs and involvement of community members have popularized the OSS movement. The major hindrance in adoption of OSS is lack of technical support for implementation[6]. Koha, NewGenLib, EverGreen, OpenBiblio, ABCD are some of Open source library management software. Joomla, Drupal, WordPress etc. are some of the open source software used for content management and GSDL, DSpace, Eprintetc are open source software used for creation of institutional repositories[7].

LITERATURE REVIEW

Few studies are available on the bibliometric assessment of global research productivity of open source software. Kumar, Ashok et al (2019)[8]examined 2911 research publications on OSS, sourced from Scopus database for the period 2009-18. The data was analysed to study yearly output, language of sources, medium of communication, prolific authors, subject wise distribution, most productive organisations. Ramesh and Jayaprakash (2016)[9] assessed the global research output of open source software using IEEE database during 1994-2010. The findings revealed low publication output of 223 articles. It showed that research is in nascent stage. Palmer & Choi (2014)[10] examined the library open source software research by descriptive literature review using databases Library Literature and Information Science Full Text and Library, Information Science and technology Abstracts (LISTA). The research was analysed on parameters like year of publication,

type and topic of article, software type etc. Ammarukleart and Kim (2017)[11] investigated the research trends in institutional repositories using bibliometric and text mining techniques for the period 2005-2015. The data was derived from the LISA and the Web of Science citation databases. The study analysed 603 articles published in 109 journals and inclusion of various research topics in literature like research data, data management, linked data etc. Ahmad et al (2018)[12] conducted a bibliometric study on the digital library for the period 2002 to 2016. The data was retrieved from ISI web of Science and analysed. The findings revealed 4206 documents with 2016 as the most productive year, Electronic Library as the top source title, the USA most productive country, Illinois University as the most productive organization and Fourie I as the most prolific author.

OBJECTIVES OF THE STUDY

The main objectives of this study is to focus on the growth rate in global publication output in OSS research and its citation impact during 1999-18; (ii) To examine the contribution and citation impact of top 10 most productive countries and top 15 most productive organizations and authors; (iii) To identify the international collaborative share of top 10 most productive countries; (iv) To analyse the global publication output by broad subject categories; (v) To study the medium of communication and to study the characteristics of highly cited papers on OSS research.

METHODOLOGY

The publications data on OSS research of the world covering 20 year period 1999-2018 was sourced from the Web of Science database. Main search strategy was formulated using significant keywords such as "open source software" OR "open source softwares" OR "OSS" OR "DSpace" OR "Koha" OR "EPrints" OR "New Gen Lib" OR "EverGreen" OR "Fedora" OR "GreenStone" OR "GSDL" and limited to subject category "Information Science & Library Science". The string was restricted to the period 1999-2018. The search string was restricted by individual name in "country tag" to retrieve OSS research output of the top 10 most productive countries. The search string was further restricted to analytical functions and tags by subject, country, source-title, author wise and organisation wise etc. Citations data was collected from date of publication till 15 October 2019.

DATA ANALYSIS AND RESULTS

Global publications data on OSS, sourced from Web of Science database, cumulated to a total of 341 publications in 20 years during 1999-18, registered 23.12% annual growth, up from 2 in 1999 to 19 publications in 2018. The absolute growth in 10 years was 170.65%, witnessing an increase from 92 in 1999-08 to 249 publications in 2009-18. The citation impact of OSS research averaged to 7.54 citations per paper during the period. Its 10 year citation impact dropped from 10.68 CPP during 1999-08 to 6.38 CPP during 2009-18 (Table 1). Of the total global publications (341), 87.98% (300) appeared as articles, 7.62% (26) as book reviews, 4.11% (14) as conference papers, 2.05% (7) as editorials, and the rest as reviews, letters and news item (0.88% to 0.29%). Of the global publication output, 91.20% (311) appeared in English, followed by 4.69% (16) in Spanish, 2.64% (9) in Hungarian, 0.59% (2) in German and 0.29% (1) each in Italian and Portuguese and the rest in other languages.

Table 1: Annual and Cumulative Growth of Publications in OSS during 1999-18

Period	TP	TC	CPP	Period	TP	TC	CPP
1999	2	9	4.5	2011	19	276	14.53
2000	5	2	0.4	2012	28	308	11
2001	5	84	16.8	2013	26	158	6.08
2002	8	65	8.13	2014	24	118	4.92
2003	5	22	4.4	2015	27	101	3.74
2004	6	71	11.83	2016	35	70	2
2005	15	171	11.4	2017	27	37	1.37
2006	12	134	11.17	2018	19	0	0
2007	13	224	17.23	1999-08	92	983	10.68
2008	21	201	9.57	2009-18	249	1589	6.38
2009	25	274	10.96	1999-18	341	2572	7.54
2010	19	247	13				

TP=Total Publications; TC=Total Citations; ACPP= Average Citations Per Paper

Top 10 Countries in OSS Research

The global research output on OSS had originated from 49 countries during 1999-18, of which 28 countries published 1-4 papers each, 14 countries 5-8 papers each, 3 countries 10-18 papers each and 4 countries 20-111 papers each. The top 10 most productive countries in OSS research are as follows: USA with highest publication share (32.55%), followed by India (10.26%), England (7.62%), China (5.87%), Spain (5.28%), Canada (3.81%), Germany (2.93%), Australia and France (2.35% each), NewZealand (2.05%) during 1999-18. The global publication share varied between 2.05% and 32.55% and together contributed 75.07% global share during 1999-18. The 10 year global

Table 2: Scientometric Profile of Top 10 Countries in OSS Research during 1999-18

Country	Total Papers			Share of Papers			TC	CPP	ICP	%ICP	RCI
	1999-08	2009-18	1999-18	1999-08	2009-18	1999-18					
USA	30	81	111	32.61	32.53	32.55	981	8.84	18	16.22	1.17
India	9	26	35	9.78	10.44	10.26	124	3.54	1	2.86	0.47
England	14	12	26	15.22	4.82	7.62	303	11.65	8	30.77	1.55
China	2	18	20	2.17	7.23	5.87	209	10.45	13	65.00	1.39
Spain	4	14	18	4.35	5.62	5.28	209	10.45	13	72.22	1.39
Canada	4	9	13	4.35	3.61	3.81	106	8.15	3	23.08	1.08
Germany	1	9	10	1.09	3.61	2.93	51	5.1	4	40.00	0.68
Australia	2	6	8	2.17	2.41	2.35	109	13.63	2	25.00	1.81
France	0	8	8	0	3.21	2.35	46	5.75	7	87.50	0.76
New Zealand	3	4	7	3.26	1.61	2.05	40	5.71	1	14.29	0.76
Total of 10 Countries	69	187	256				2178	8.51	70	27.34	1.13
World	92	249	341				2572	7.54			1
Share of top 10 countries in global output	75	75.1	75.07								

TP=Total Publications; TC=Total Citations; CPP= Citations Per Paper;
ICP= International Collaborative Papers; RCI= Relative Citation Index

publication share in OSS research registered an increase by 5.06% in China followed by 3.21% in France, 2.52% in Germany, 1.27% in Spain, 0.66% in India and 0.24% in Australia and dropped in countries i.e. 10.4% in England, 0.74% in Canada and 0.08% in USA. Six countries registered relative citation index above the world average of 1: Australia (1.81), England (1.55), China and Spain (1.39 each), USA (1.117) and Canada (1.08) during 1999-18.

International Collaboration

The international collaborative publications (ICP) share of top 10 countries in OSS research ranged from 2.86% to 87.50%, with France accounting for the highest ICP share (87.50%), followed by Spain (72.22%), China (65%), Germany (40%), England (30.77%), Australia (25%), Canada (23.08%), USA (16.22%), NewZealand (14.29%) and India (2.86%) during 1999-18. (Table 2)

Subject-Wise Distribution of Research Output

The global OSS research published during 1999-18 is distributed across six subjects (as identified in Web of Science database classification). Information Science Library Science accounted for the highest publication share (100%), followed by computer science information systems (30.79%), management (10.26%), communication (3.52%), computer science interdisciplinary applications (1.76%) and history (0.59%) during 1999-18. The research output dynamism is reflected by activity index (world average activity index of a subject is taken as 100). The activity index witnessed increase (from 1999-08 to 2009-18) in computer science interdisciplinary applications and history (from 0.00 to 138.0), management (from 62.23 to 114.34), communication (from 90.75 to 103.5) and information science library science (from 97.94 to 100.77) and decline in computer science information science (from 127.91 to 89.37). Management sub-field registered the highest citation impact per paper (CPP) of 24, followed by computer science information system (15.31), information science library science (7.54), computer science interdisciplinary applications (6), history (2.5), communication (2.17) (Table 3).

Table 3: Subject Wise Breakup of Global Publications Output

Subject	Total Papers			Activity Index		TC	CPP	%TP
	1999-08	2009-18	1999-18	1999-08	2009-18			
Information Science Library Science	92	249	341	97.94	100.77	2572	7.54	100
Computer Science Information Systems	37	68	105	127.91	89.37	1608	15.31	30.79
Management	6	29	35	62.23	114.34	840	24.00	10.26
Communication	3	9	12	90.75	103.5	26	2.17	3.52
Computer Science Interdisciplinary Applications	0	6	6	0.00	138	36	6.00	1.76
History	0	2	2	0.00	138	5	2.50	0.59
Total of the World	92	249	341			2572	7.54	

TP=Total Publications; TC=Total Citations; CPP= Citations Per Paper

Top 15 Organisations

Top 15 most productive global organizations in OSS research contributed 3 to 6 publications each and together they contributed 66 publications (19.35% global share) and received 486 citations (18.9% global share) during 1999-18.On further analysis, it was observed:

- Seven organisations registered publications productivity above the group average of 4.4: University of Kentucky, USA (6), City University of Hong Kong, Indiana University, USA, Texas Tech University, USA, University of Arizona, USA, University of Illinois System, USA and Victoria University Wellington, N.Z. (5 papers each).
- Seven organisations registered citation impact per paper and relative citation index above the average impact of 7.42 and 0.98: University of Granada, Spain (19.0 and 2.52), City University of Hong Kong (18.8 and 2.49), Nanyang Technological University, National Institute of Education, Singapore (11.0 and 1.46), University of Illinois USA (10.0 and 1.33), Texas Tech University, USA (9.4 and 1.25), Indian Statistical Institute, Bangalore (9.0 and 1.19) and Pennsylvania Commonwealth System of Higher Education, USA (8.5 and 1.13) (Table 4)

Table 4: Scientometric Profile of Top 15 Most Productive Organisations in OSS during 1999-18

Sl. No.	Name of Organisation	TP	TC	CPP	HI	RCI
1	University of Kentucky, USA	6	22	3.67	3	0.49
2	City University of Hong Kong	5	94	18.8	5	2.49
3	Indiana University, USA	5	37	7.40	2	0.98
4	Texas Tech University, USA	5	47	9.40	2	1.25
5	University of Arizona, USA	5	8	1.60	2	0.21
6	University of Illinois USA	5	50	10.0	2	1.33
7	Victoria University Wellington, N. Z.	5	7	1.4.0	1	0.19
8	Nanyang Technological University, National Institute of Education, Singapore	4	44	11.00	2	1.46
9	Pennsylvania Commonwealth System of Higher Education, USA	4	34	8.5	2	1.13
10	Texas AM University, USA	4	9	2.25	2	0.3
11	University of Granada, Spain	4	76	19.00	2	2.52
12	University of Tennessee, USA	4	7	1.75	1	0.23
13	University of Wisconsin, USA	4	6	1.50	1	0.2
14	California State University, USA	3	18	6.00	2	0.8
15	Indian Statistical Institute, Bangalore	3	27	9	2	1.19
	Total of 15 Organizations	66	486	7.42	2.07	0.98
	Global Total	341	2572	7.54		1
	Share of 15 organizations in Global Total	19.35	18.9			

TP=Total Publications; TC=Total Citations; CPP= Citations Per Paper; HI= H-Index; RCI= Relative Citation Index

Top 15 Authors

The research productivity of top 15 most productive authors in OSS research varied from 2 to 5 publications and together they contributed 12.02% (41) global publication share and 11.66% (300) global citation share during 1999-18.

- Seven authors registered publication productivity above the group average of 2.7 per author: N. Choi (5 papers), B. Chawner and V. Singh (4 papers each), H.L. Chen, E. De Smet, O. Nov and Y.L. Fang (3 papers each);
- Four authors registered citation impact per paper and relative citation index above the group average of 7.32 and 0.97: O. Nov (34.33 and 4.55), X.G. Chen (20.5 and 2.72), D. Bainbridge (16.5 and 2.19), G. Bissels (13.5 and 1.79) during 1999-18.

Table 5: Scientometric Profile of 15 Most Productive Authors in OSS during 1999-18

Sl. No.	Author	Affiliation	TP	TC	CPP	HI	RCI
1	N. Choi	University of Kentucky, USA	5	20	4.0	3	0.53
2	B.Chawner	Victoria University Wellington, N.Z.	4	0	0.0	0	0.00
3	V. Singh	University of Tennessee, Knoxville, USA	4	11	2.75	2	0.36
4	H.L. Chen	Long ISI University Greenvale, USA	3	12	4.0	2	0.53
5	E. De Smet	University of Antwerp, Belgium	3	13	4.33	2	0.57
6	O. Nov	New York University, Brooklyn, USA	3	103	34.33	2	4.55
7	Y.L. Fang	City University of Hong Kong	3	8	2.67	1	0.35
8	S.M.Z. Ahmed	University of Dhaka, Bangladesh	2	2	1.0	1	0.13
9	B. Albee	EBSCO Information Services, USA	2	11	5.5	2	0.73
10	K.T.Anuradha	Indian Institute of Science, Bangalore	2	10	5.0	2	0.66
11	D. Bainbridge	University of Waikato, N.Z.	2	33	16.5	2	2.19
12	G.Bissels	University of Applied Sciences Grisons, Switzerland	2	27	13.5	2	1.79
13	N.C. Chang	Tatung University, Taiwan	2	3	1.5	1	0.2
14	X.G. Chen	Shanghai Jiao Tong University, China	2	41	20.5	2	2.72
15	S.Cherukodan	Cochin University of Science &Technology, India	2	6	3.0	1	0.4
	Total of 15 Authors		41	300	7.32	1.67	0.97
	Total of World		341	2572	7.54		1
	Share of 15 authors in global output		12.02	11.66			

TP=Total Publications; TC=Total Citations; CPP= Average Citations Per Paper; HI= H-Index; RCI= Relative Citation Index

Medium of Research Communication

Eighty seven per cent (87%) of the total world output on OSS research appeared in journals. The top 15 most productive journals contributed 5-29 papers each and together accounted for 61% (208) of total publication output in journal medium during 1999-18. The publication share of 15 most productive journals decreased from 71.74% to 57.03% between 1999-08 and 2009-18. *Program Electronic Library and Information Systems* (29 papers) was the most productive journal, followed by *Electronic Library* (27 papers), *Library Hi Tech* (25 papers) and Library *Journal* (24 papers) during 1999-18 (Table 6)

Table 6: Distribution of Journal Papers by Serial Productivity in OSS during 1999-18

Sl. No.	Name of Journal	No. of Papers		
		1999-08	2009-18	1999-18
1	Program Electronic Library and Information Systems	13	16	29
2	Electronic Library	4	23	27
3	Library HiTech	4	21	25
4	Library Journal	14	10	24
5	Information Technology and Libraries	6	10	16
6	Journal of the American Society For Information Science and Technology	4	9	13
7	Information Management	3	9	12
8	International Journal of Information Management	4	8	12
9	Online Information review	4	8	12
10	Journal of Information Technology	3	7	10
11	DESIDOC Journal of Library Information and Technology	0	6	6
12	Library Trends	4	2	6
13	Scientometrics	0	6	6
14	Digital Library Perspectives	0	5	5
15	Information Research. An International Electronic Journal	3	2	5

SUMMARY AND CONCLUSION

Research publications on OSS research sourced from Web of science database was analysed to provide a quantitative and qualitative description of world research output in 20 years (1999-18). The study conducted across 49 countries registered 23.12% growth (341 publications) and citation impact averaged to 7.54 citations per paper during 1999-18, which decreased from 10.68 during 1999-08 to 6.38 during 2009-18. 75.07% global OSS research output came from top 10 countries in OSS research. USA accounted for the highest publication share (32.55%), followed by India (10.26%), England (7.62%), China (5.87%), Spain (5.28%) and Canada, Germany, Australia, France, New Zealand (3.81% to 2.05%) during 1999-18'. Six countries registering relative citation index above the world average are: Australia (1.81), England (1.55), China and Spain (1.39 each), USA (1.17) and Canada (1.08) during the period. The ICP share of top ten countries varied from 2.86% to 87.50% of their national productivity. Information Science Library Science is the most sought after subject in OSS research accounting for 100% publication share, followed by Computer Science Information Systems (30.79%), Management (10.26%) etc. during the period. The Top 15 most productive organizations and authors together contributed 19.35% and 12.02% global publication share and 18.9% and 11.66% global citation share respectively during 1999-18. Journals medium accounted for 87% global share in OSS research. The top 15 most productive journals accounted for 61% of totalpublication output in journal medium during 1999-18. *Program Electronic Library and Information Systems* **c**ontributed maximum papers (29), followed by *Electronic Library* (27) and*Library HiTech* (25) during 1999-18. The low productivity of OSS research in 20 years depict that OSS research is in its nascent stage. In order to catalyse the research of open source software, collaboration at national and international level is essential. Open source software serve as a boon for

libraries with low budget. It has helped to marginalise the digital divide. Libraries are able to optimize their resources with the use of open source software. It is imperative to boost the OSS research.

REFERENCES

1. Singh Negi, D. Open Source Software using New GENLIB: A case study of international management institute Bhubaneswar. Library Hi Tech News, 2014, 31(9), 9-10. https://doi.org/10.1108/LHTN-07-2014-0056 (Accessed on 10 Dec 2019).
2. Salve, Anand, Lihitkar, Shalini R and Lihitkar, Ramdas. Open source software as tools for libraries: An overview. DESIDOC Journal of Library and Information Technology, 2012. 32(5), 381-387.
3. Pruett, J, Choi, N. A comparison between select open source and proprietary integrated library systems, Library Hi Tech 2013, 31(3), 435-454.https://doi.org/10.1108/LHT-01-2013-0003 (Accessed on 10 Dec 2019).
4. Singh, M and Sanaman, G. Open Source integrated library management systems: Comparative analysis of Koha and New GenLib. The Electronic Library, 2012, 30(6), 809-832. https://doi.org/10.1108/LHT-01-2013-0003 (Accessed on 10 Dec 2019).
5. Kamper, R and Kaushik, P. Economics of open source library software: Evidences from Indian libraries. Global Knowledge, Memory and Communication 2019, 68(4-5), 337-355 https://doi.org/10.1108/GKMC-04-2018-0036 (Accessed on 10 Dec 2019).
6. Adera Amolla, B. Feasibility of adaptation of open source ILS for libraries in Kenya: A practical evaluation. The Electronic Library, 2013, 31(5), 608-634.https://doi.org/ 10.1108/EL-12-2011-0171 (Accessed on 10 Dec 2019).
7. Londhe, Nagesh L. and Patil, Suresh K. Open source library management systems: A survey and present developmental status. International Journal of Library and Information Science, 2015, 4(1), 38-54.
8. Kumar, Ashok, Kumar, Sunil and Chanchal. Analysis of global publications on open source software in libraries. In. Proc. Of the international conference: 3rd DLA-SRFLIS summit 2019, 394-399.
9. Ramesh, K and Jaya Prakash M. Scientometric analysis of open source library software. Indian Journal of Information Sources and Services, 2016, 6(2), 27-29.
10. Palmer, A. and Choi, N. The current state of library open source software research: A descriptive literature review and classification. Library Hi Tech, 2014, 32(1), 11-27. https://doi.org/10.1108/LHT-05-2013-0056 (Accessed on 10 Dec 2019).
11. Ammarukleart, S and Kim, J. Institutional repository research 2005-2015: A trend analysis using bibliometrics and text mining. Digital Library Perspectives, 2017, 33(3), 264-278.https://doi.org/10.1108/DLP-07-2016-0027 (Accessed on 10 Dec 2019).
12. Ahmad, K, Jian Ming, Z. and Rafi, M. Assessing the digital library research output: Bibliometric analysis from 2002 to 2016. The Electronic Library, 2018, 36(4), 696-704. https://doi.org/10.1108/EL-02-2017-0036 (Accessed on 10 Dec 2019).

❖❖❖