

University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

Library Philosophy and Practice (e-journal)

Libraries at University of Nebraska-Lincoln

June 2021

Bibliometrics Analysis and Comparison of Global Research Literatures on Research Data Management extracted from Scopus and Web of Science during 2000 - 2019

Pallab Pradhan

Information and Library Network Centre (INFLIBNET), Gandhinagar, ppradhan86@gmail.com

Lavji N. Zala

Sardar Patel University, Vallabh Vidyanagar, lavji_zala@rediffmail.com

Follow this and additional works at: <https://digitalcommons.unl.edu/libphilprac>



Part of the [Library and Information Science Commons](#)

Pradhan, Pallab and Zala, Lavji N., "Bibliometrics Analysis and Comparison of Global Research Literatures on Research Data Management extracted from Scopus and Web of Science during 2000 - 2019" (2021).

Library Philosophy and Practice (e-journal). 5519.

<https://digitalcommons.unl.edu/libphilprac/5519>

Bibliometrics Analysis and Comparison of Global Research Literatures on Research Data Management extracted from Scopus and Web of Science during 2000 - 2019

Pallab Pradhan, Scientist-C (Library Science), Information and Library Network
(INFLIBNET) Centre, Gandhinagar - 382007, Gujarat (INDIA)

Dr. Lavji N Zala, Assistant Professor, Department of Library & Information Science, Sardar
Patel University, Vallabh Vidyanagar - 388120, Gujarat (INDIA)

Abstract: Researchers, to conduct any bibliometric analysis prefer to retrieve publications data mostly from Elsevier's Scopus or/and Clarivate Analytics' Web of Science (WoS) databases, though many other platforms/databases, i.e. Google Scholar, Dimensions, Crossref, PubMed, etc. are now available those are providing bibliographic data of publications. This study is based on the globally published literatures on research data management during 2000 – 2019 (20 years of duration) data extracted from the Scopus & Web of Science (WoS) databases and their Merged file. The analysis and results compares the similarity and differences in between Scopus & WoS, and further each one of them with the Merged file. The study reveals that around 32% of globally published literatures on research data management were indexed in both the Scopus and WoS databases. It compares both the sources in terms of parameters like annual literatures growth & trends, top authors production, authorship & collaboration pattern, most relevant sources & affiliations, country scientific production and international collaboration, etc. along with the merged file of both the datasets as well wherever possible.

Keywords: Bibliometric Analysis; Research Data Management; Scopus; Web of Science; BiblioShiny; Bibliometrix

1. Introduction

Bibliometrics is the application of mathematics and statistical methods to books and other media of communication. While coining the term Bibliometrics, Alan Pitchard hoped that this term Bibliometrics will be used explicitly in all studies which seek to quantify the processes of written communication and will quickly gain acceptance in the field of information science (Pitchard, 1969), and that truly happened in due course of time. It's not only became popular in the field of information science, now it is being used in almost all fields of studies to conduct different kinds of bibliometric analysis on the body of literatures across fields/disciplines. It can be said that Bibliometrics itself is a field of study now with in-depth theories, methods and its applications.

Bibliometricians and researchers to conduct any bibliometric analyses were preferring to retrieve publications data mostly from Institute for Scientific Information founded in 1956 in Philadelphia by Eugene Garfield, later, which became Thomson Reuters - ISI database in 1992 (Currently, it is Clarivate Analytics' Web of Science database) since last few decades. Then Google's Google Scholar and Elsevier's Scopus database came into market in 2004. Scopus gradually become more popular, and huge competitor for ISI-Thomson Reuters, now Web of Science database. Even, Elsevier proudly declared Scopus as the world's largest scientific abstract database as well. Currently, there are many other platforms/databases which are providing bibliographic data of publications, i.e. Google Scholar, Dimensions, Crossref,

PubMed, etc. However, each database differs from each other in terms of their scope and coverage. And, to prove the same, many studies have been conducted in the past to evaluate and compare each database's scope, coverage and limitations. (Sánchez et al., 2017) compared the WoS and Scopus databases for conducting bibliometric analyses and suggested no superiority of one over the other. (Fernández, Barbosa, & Guerrero, 2010) argued that WoS and Scopus are complementary to each other based on their study of chemical literatures retrieved. Thus, it is not a wise idea to argue and state that which database is the best to be used for bibliometrics analyses. It solely depends on the users/researchers to choose the database with which they want to conduct or carry out such studies. And, it is always best to use data from both WoS & Scopus, or from all the databases available if possible.

The authors provide some reasons why most of the Bibliometricians or researchers prefer data from only a single database to conduct their study, or why most of the bibliometric studies are based on publication records or bibliographic data retrieved only from one single database. It's quite obvious that it's not always possible for each and every researcher to get access to both the WoS and Scopus databases at once at their institutions. Not each and every researcher's affiliated Institutions may have a library subscription to access both the databases at the same time due to funding issue or budget crunch for library resources subscriptions as both are commercial and much expensive. Also, it's quite difficult and time consuming to merge the publication records or bibliographic data retrieved from both the databases to conduct a study due to their variance in data formats/standards and field tags used. If that's done so, it's not wise to argue and be assured that the merged data would be correct & perfect, so as its results. This is because the authors' affiliations, address details differ at large while indexing, and not at all remain same in different databases. And, if any researcher is conducting a bibliometric study focused on parameters like top affiliations, country scientific production, country and institutes collaboration, etc. then the results derived may not be correct due to the mentioned variation issues in the data. The issue occurs most while merging and cleaning up duplicate records from the data extracted from different databases manually thorough MS Excel or any other software/tools, and it's difficult to choose which database's data to be kept, and that impact the data accuracy and results later on.

The authors present a detailed bibliometric study of global research literatures on research data management retrieved from Scopus and Web of Science during 2000 – 2019 through this article. They compare both the databases in terms of their differences on different bibliometric parameters based on retrieved global research literatures on research data management. Also, they merge the data from both the databases and provide a detailed study on few parameters wherever possible. Attempts have been made to analyse and explore the annual literature growth & trends, authorship pattern, author's collaboration network & top productive authors, top keywords used based on their frequency of occurrences & keywords clustering, most relevant sources, etc. The authors conduct three separate bibliometric analyses using Scopus, WoS, and Merged data of both to demonstrate differences among results between Scopus, WoS, and Merged data. This paper makes major contribution to the literature based on the bibliometric analysis of global research literatures on research data management.

2. Related Works

(Bakkalbasi, Bauer, Glover, & Wang, 2006) reported a comparison study between Google Scholar, Scopus, and Web of Science using citation analysis from the selected eleven journal

titles with varying impact factors from two disciplines (oncology and condensed matter physics) using the Journal Citation Reports (JCR). All articles published in the selected titles were retrieved for the years 1993 and 2003, and a stratified random sample of articles was chosen, resulting in four sets of articles. The citation counts for each research article were extracted from these three sources during the week of November 7-12, 2005. Further, the actual citing references for a subset of the articles published in 2003 were also gathered from each of the three sources. Their study did not identify any one of these three resources as the answer to all citation tracking needs. Scopus showed strength in providing citing literature for current (2003) oncology articles, while Web of Science produced more citing material for 2003 and 1993 condensed matter physics, and 1993 oncology articles. All three tools returned some unique material. (Gavel & Iselid, 2008) presented the overlaps calculation in between the journal lists of Web of Science and Scopus databases. From the study it was found that the number of titles covered by Scopus exceeds the number of titles covered by WoS. Further, they revealed that the superior coverage of Scopus is mainly associated with the science, technical & medical (STM) area, where WoS has comparatively few unique titles in the field. (Fernández, Barbosa, & Guerrero, 2010) carried out a comparison study in between the two most extended platforms of scientific information: Web of Science and Scopus, applying quantitative methods on certain parameters like literature growth pattern; the overlapping among the two databases, the dispersion of the articles in the journals, concentration measures, and possible correlations, etc. They searched in the area of the chemical engineering in both databases between 1999 and 2006. The results showed the existence of a high likeness between Web of Science and Scopus, turning out complementary but not exclusive, regarding their possible use for the chemical engineers. (Chirici, 2012) assessed the scientific productivity of Italian forest researchers using the Web of Science, SCOPUS and SCIMAGO databases. In his study, he compared the WoS and SCOPUS databases with respect to three indexes (number of publications, number of citations, h-index) of the scientific productivity for university forest researchers in Italy. He opined that both WoS and SCOPUS databases were suitable sources of information for evaluating the scientific productivity of Italian authors, and the two databases did not produce meaningful differences for any of the three indexes mentioned. (Archambault, Campbell, & Larivière, 2013) used macro level bibliometric indicators to compare results obtained from the WoS and Scopus database. Their study showed extremely high correlations ($R^2 \approx .99$) between the measures obtained with both databases for the number of papers and the number of citations received by countries. The paper provided evidence that indicators of scientific production and citations at the country level are stable and largely independent of the database. (Wagner, 2015) practically compared both the Scopus and Web of Science Core Collection based on their features, scope and coverage, etc. He found out that the Scopus had somewhat stronger international/non-English coverage as compare to WoS, and WoS appeared to be catching up. Scopus had much stronger coverage in fields like Social Sciences, Arts & Humanities. (Mongeon & Paul-Hus, 2016) described the journal coverage of both Web of Science and Scopus and assessed whether some field, publishing country and language were over or underrepresented in those databases. They compared the coverage of active scholarly journals in WoS (13,605 journals) and Scopus (20,346 journals) with Ulrich's extensive periodical directory (63,013 journals). Their results indicated that the use of either WoS or Scopus for research evaluation may introduce biases that favour Natural Sciences and Engineering as well as Biomedical Research to the detriment of Social Sciences and Arts and Humanities. Similarly, English-language journals are overrepresented to the detriment of other

languages. Further, it was found that both databases shared these biases, their coverage differs substantially, and as a consequence, the results of bibliometric analyses may vary depending on the database used. Thus, WoS and Scopus should be used with caution, especially when comparing different fields, institutions, countries or languages. (Sánchez, de la Cruz Del Río Rama, & García, 2017) showed the current state of scientific research regarding wine tourism, by comparing the platforms of scientific information WoS and Scopus applying quantitative methods. A bibliometric study of the publications indexed in WoS and Scopus was conducted with a set of 238 articles and 122 different journals obtained. They analysed the correlation between increases, coverage, overlap, dispersion and concentration of documents. Based on the results of the comparative study, they concluded that WoS and Scopus databases differ in scope, data volume and coverage policies with a high degree of unique sources and articles, resulting both of them complementary and not mutually exclusive. Scopus covers the area of wine tourism better, by including a greater number of journals, papers and signatures. (Echchakoui, 2020) conducted a bibliometric study by retrieving papers on sales force literature from Scopus and WoS databases covered from 1912 to 2019, further he merged both the datasets as well and compared. The results showed that there were many disparities between WoS and merged database, and between Scopus and merged database regarding bibliometric analyses, especially among primary productive authors, the most influential papers, and keyword occurrences. His research proposed a four-step procedure that merges these two databases to allow more reliable bibliometric analyses.

3. Objectives

The main objectives of this study is to conduct a bibliometric analysis of the global research literatures on research data management extracted from databases, i.e. Scopus and Web of Science published during 2000-2019, and compare both the data files in terms of parameters like annual literatures growth & trends, top authors production, authorship & collaboration pattern, most relevant sources & affiliations, country scientific production and international collaboration, etc. along with the merged file of both the datasets as well wherever possible.

4. Methodology

4.1 Sources of Data

For the collection of global research literatures data on research data management for a span of 20 years during 2000-2019, the Scopus from Elsevier and Web of Science Core Collection from Clarivate Analytics databases were accessed.

4.2 Data Collection

The study is focused to the global research literatures on research data management” published during 2000-2019 extracted from Scopus and Web of Science Core Collection databases. A total of 526 and 368 publications were retrieved from Scopus and WoS database respectively limiting to the types of document to only Article, Conference Paper/Proceedings Paper, Book Chapter, Review, and Editorial. The data was collected on 23rd December 2019 / Wednesday. All types of published documents were considered for this study. The following advanced search queries were used specifically in each database to retrieve the required:

WoS: TS=("Research Data Management")

Scopus: TITLE-ABS-KEY ("Research Data Management")

Here, the authors found that from the total of 526 Scopus extracted data, 07 titles were published twice in different journals hence not considered as duplicates. Further, from the total of 894 merged data file (526+368=894), 293 (32.77%) publications were found duplicates in titles & DOIs, hence discarded keeping 01 title from each set of duplicates, the total came up to 601 publications. Thus, the total of 601 publications were considered for this study as merged data file.

In the Web of Science extracted file, few documents were assigned as both “Article; Proceedings Paper” and “Article; Book Chapter”, but those have been considered only as ‘Proceedings Paper’ and ‘Book Chapter’ respectively. This happens because the Proceedings Paper and Book Chapters are generally research articles published in Conference Proceedings or Volumes, hence assigned as both in Web of Science. Also, both the Proceedings Paper and Conference Paper have been recoded as Conference Proceedings in the merged file. Counting of Index Keywords have not been done in the merged file as it was not required for the study.

4.3 Data Analysis

For this study, for duplication checking and merging of extracted datasets, MS Excel was used. And, to perform the required quantitative analysis along with their visualisation on the collected literature data from Scopus and WoS, the BiblioShiny app of “Bibliometrix”, R package was extensively used. Further, the collaboration networks of authors and countries, and keywords clustering were produced using the same tool.

For merging of both the Scopus and WoS collected datasets, both the raw extracted files should be uploaded into the BiblioShiny app where the user would get the output as exported file which can be saved in excel format for later use. If the user would compare it can be seen that its assigned CODEs are almost similar in both the exported files which help the user to merge both. If the datasets are having less number of records, first manual duplication, and then merging can be done as done for this study.

5. Findings and Discussions

5.1. Document Types

From the data collected from Scopus, it was found that out of total 526 published literatures globally on research data management, 292 (55.5%) were articles, followed by 194 Conference Paper/Conference Proceedings. Whereas data collected from the WoS, 206 (55.98%) were articles, followed by 132 (35.87) Conference Paper/Conference Proceedings. As far the Merged data file is concerned, 343 (57.07%) were articles, followed by 211 (35.11%) Conference Paper/Conference Proceedings. The Table-1 shows the document types with the number of published literatures globally on research data management during the studied period.

Table-1: Document Types with the number of Published Literatures

	Scopus	WoS	Merged
Document Type	Documents	Documents	Documents
Article	292	206	343
Conference Paper/Conference Proceedings	194	132	211
Book Chapter	19	10	25

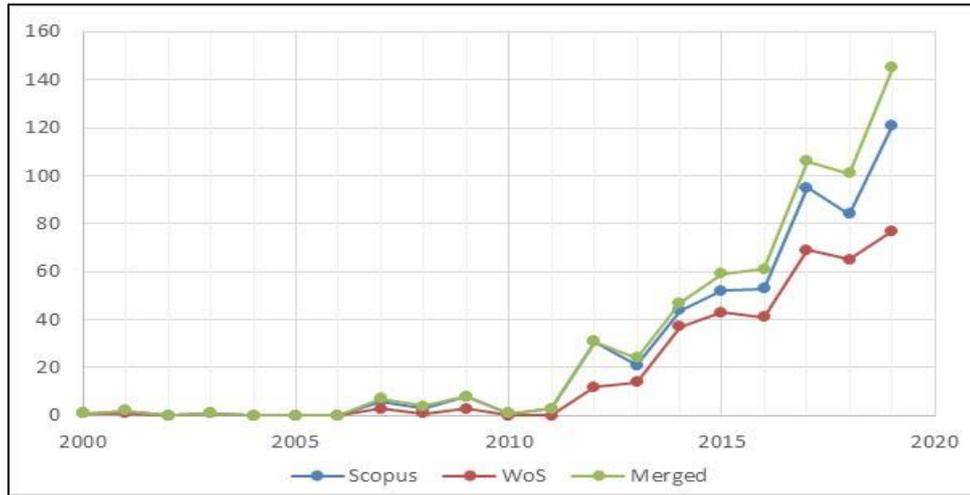
Review	19	13	17
Editorial	2	7	5
Total	526	368	601

5.2. Literature Growth/Trends

The Table-2 shows the year-wise literatures published globally on research data management during the study period 2000 - 2019. The annual percentage publications growth rates (Annual Growth rate) were, i.e. for the data collected from Scopus were 37.67%, for the data collected from WoS were 39.67%, and for the merged data file it were 39.34% respectively. Graph-1 shows the publication growth trends of published literatures globally on research data management during 2000 - 2019 for Scopus, WoS, and merged file. It can be seen that literature on research data management in both Scopus and WoS started to get published from same 2000. But, there was no literatures published in both the databases from 2004 to 2006. There was in consistency in literature growth trends in both the databases. However, it can be seen from the Table-2 that there was consistent growth in literatures published in the Merged data file from 2011 onwards with a slight decline in 2018.

Table-2: Year-wise Literatures Published Globally

	Scopus	WoS	Merged
Year	Documents	Documents	Documents
2000	1	1	1
2001	2	1	2
2002	0	0	0
2003	1	1	1
2004	0	0	0
2005	0	0	0
2006	0	0	0
2007	6	3	7
2008	3	1	4
2009	8	3	8
2010	1	0	1
2011	3	0	3
2012	31	12	31
2013	21	14	24
2014	44	37	47
2015	52	43	59
2016	53	41	61
2017	95	69	106
2018	84	65	101
2019	121	77	145
Total	526	368	601



Graph-1: Annual Literature Growth Trends

5.3. Authorship and Collaboration Pattern

The Table-3 shows the main information on published literatures on Authors aspect, i.e. Authors & Authors Collaboration as identified in Scopus, WOS, and merged data file. The annual percentage publications growth rates (Annual Growth rate) were, i.e. for the data collected from Scopus was 37.67%, for the data collected from WoS was 39.67%, and for the Merged data file it was 39.34% respectively. The number of authors per document was 2.71 in Scopus versus 2.76 in WoS which was slightly greater in WoS, and co-authors per document was 3.5 in Scopus versus 3.32 in WoS which was slightly greater in Scopus.

Table-3: Main Information on Published Literatures on Authors aspect

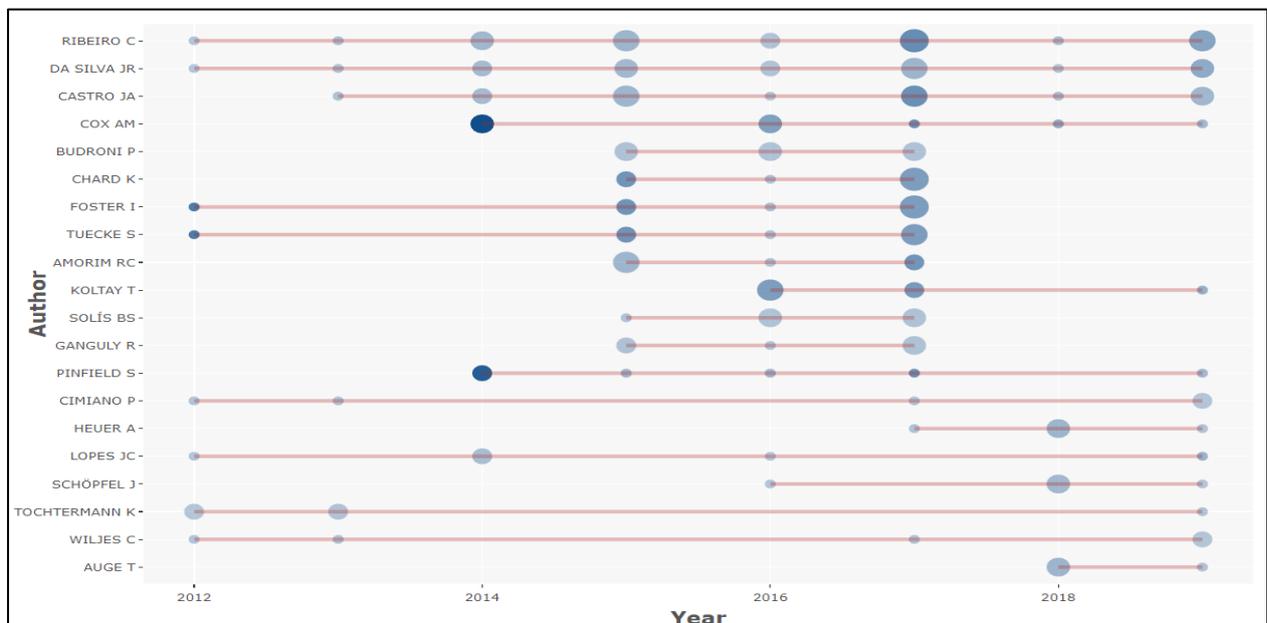
	Scopus	WoS	Merged
Authors			
Authors (Unique)	1424	1016	1543
Author Appearances	1843	1222	1988
Authors of single-authored documents	80	58	94
Authors of multi-authored documents	1344	958	1449
Authors Collaboration			
Single-authored documents	88	65	106
Documents per Author	0.369	0.362	0.39
Authors per Document	2.71	2.76	2.57
Co-Authors per Documents	3.5	3.32	3.31
Collaboration Index	3.07	3.16	2.93

The Table-4 represents the top 20 productive authors published both in Scopus and WoS based on number of publications (NP). Further, h-index for the respective authors have been shown in the same table. The author named Ribeiro C is on 1st rank with 21 number of publications (NP) in Scopus whereas he is on 2nd rank in WoS with 13 NP with an h-index of 6 and 3 respectively. The author named Da Silva JR is on 2nd rank with 17 NP in Scopus whereas he is on 1st rank in WoS with 15 NP with an h-index of 6 and 3 respectively. The author named Castro JA is on 3rd rank with 16 NP in Scopus whereas he is on 4th rank in WoS with 8 NP with an h-index of 6 and 3 respectively. The author named Cox AM is having top rank in terms of h-index in both the databases with 8 NP in each database.

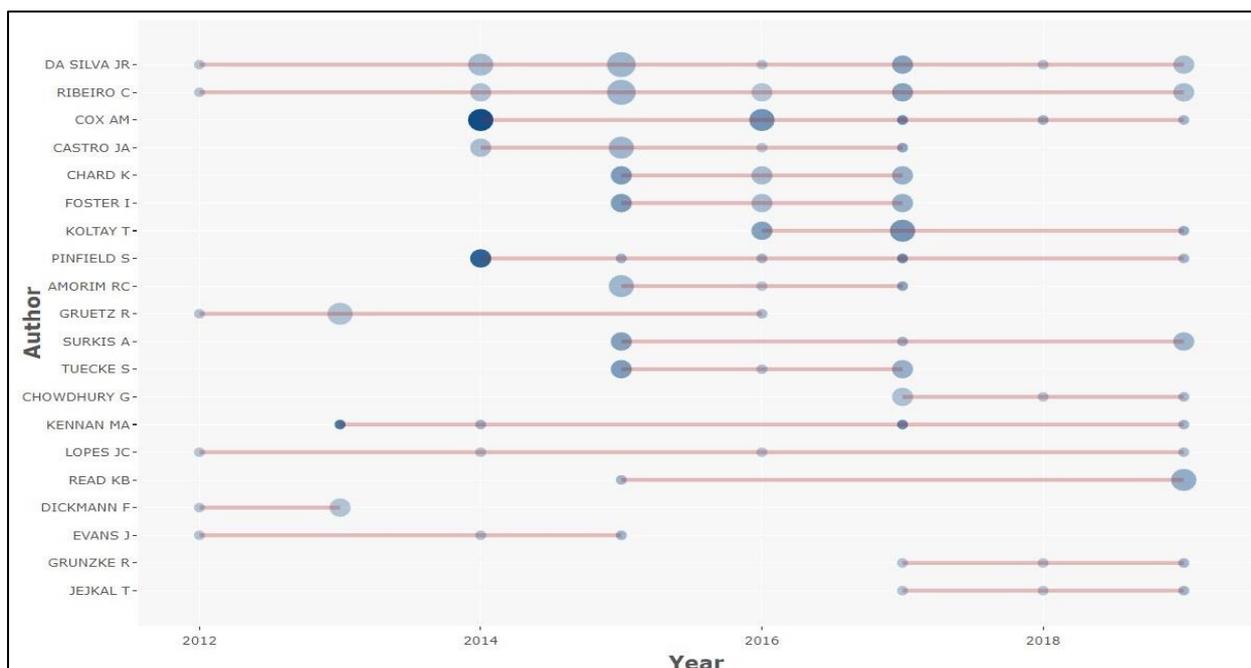
Table-4: Top 20 Productive Authors in Scopus and WoS

Author	Scopus		Author	WoS	
	h_index	NP		h_index	NP
Ribeiro C	6	21	Da Silva JR	3	15
Da Silva JR	6	17	Ribeiro C	3	13
Castro JA	6	16	Cox AM	8	9
Cox AM	8	9	Castro JA	3	8
Budroni P	2	8	Amorim RC	2	6
Chard K	5	8	Chard K	3	6
Foster I	5	8	Foster I	3	6
Tuecke S	5	8	Koltay T	4	6
Amorim RC	3	7	Pinfield S	5	6
Koltay T	4	7	Lopes JC	2	5
Solís BS	2	7	Read KB	3	5
Ganguly R	2	6	Surkis A	3	5
Pinfield S	5	6	Tuecke S	3	5
Cimiano P	2	5	Chowdhury G	2	4
Heuer A	2	5	Grutz R	2	4
Lopes JC	3	5	Kennan MA	4	4
Schöpfel J	1	5	Towe M	2	4
Tochtermann K	1	5	Dickmann F	1	3
Wiljes C	2	5	Evans J	3	3
Auge T	2	4	Grunzke R	1	3

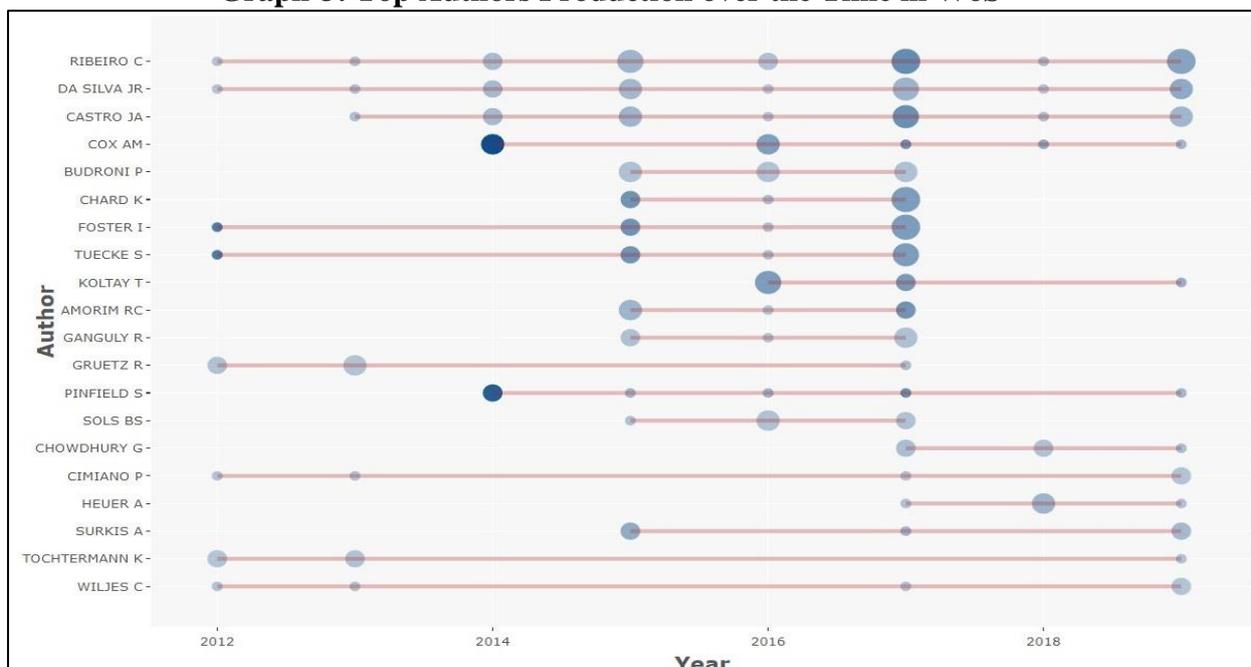
The three Graphs from Graph-2 to 4 (from top to down) represent the top 20 authors and their production over the time in Scopus, WoS, and Merged data file respectively.



Graph-2: Top Authors Production over the Time in Scopus



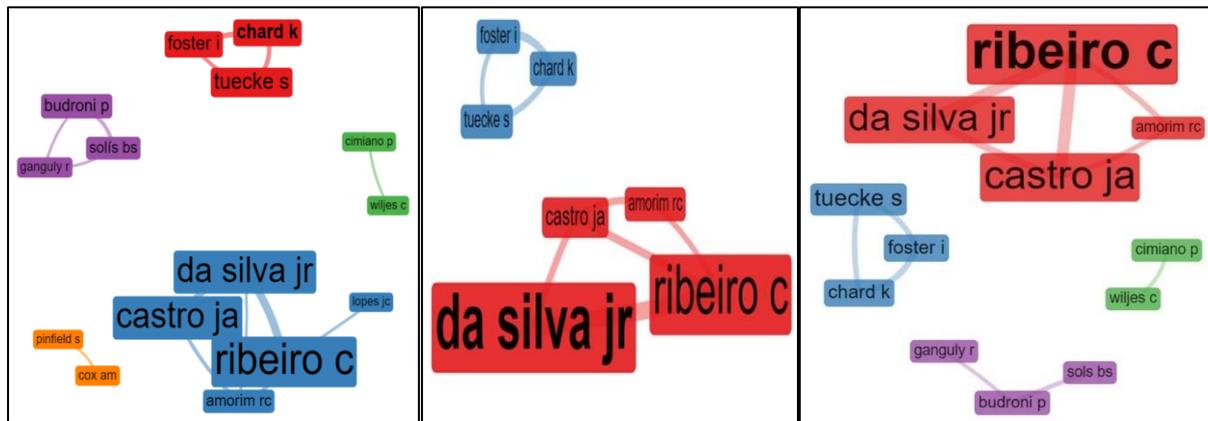
Graph-3: Top Authors Production over the Time in WoS



Graph-4: Top Authors Production over the Time in Merged data file

Graph-5 to 7 represent the author's collaboration network by two or more authors for publication in Scopus, WoS, & Merged file respectively from left to right side as shown. There were 1424 unique authors for 526 published literatures in Scopus, 1016 authors for 368 published literatures in WoS, and 1543 authors for 601 published literatures in Merged file. From the total of 526 published literatures in Scopus, 438 publications (83.27%) were multi-authored versus from the total of 368 published literatures in WoS, 303 publications (82.34%) were multi-authored, and from the total of 601 published literatures in Merged file, 495 publications (82.36%) were multi-authored. The collaboration index was also slightly greater in WoS 3.16 compare to Scopus 3.07. But, for the Merged file it was 2.93 which was below than both the databases.

The collaboration network was generated by normalizing the association between authors using the Edge Betweenness clustering algorithm where minimum edges between nodes (Authors) was considered as 5. Also, the isolated nodes were removed, thus were not considered for generating the network.



Graph-5 to 7: Authors' Collaboration Network (Scopus, WoS, & Merged file)

5.4. Top Keywords Co-occurrence and their Network

There were a total 2985 keywords [Author's Keywords (1197), and Index keywords (1788)] in Scopus versus a total of 1170 keywords [Author's Keywords (905), and Index keywords (265)] in WoS. And, there were a total 3258 keywords [Author's Keywords (1288), and Index keywords (1970)] in Merged file. However, the analysis for the most relevant keywords based on their occurrence & their clustering, and trending topics are done by using only the Author's keywords. The total of Author's Keywords and Index Keywords in Scopus was more than double as compare to WoS.

The Table-5 shows the list of the top 10 most relevant, frequently used keywords based on their occurrence/frequencies in the three data files for literatures on research data management published globally during 2000 - 2019 from left to right side as shown. The analysis was done using the Author's keywords only, excluding the Index keywords. The most frequent keyword was "research data management" which has topped the list in all data files, i.e. in Scopus with 265 occurrences, in WoS with 148 occurrences, and in Merged file with 289 occurrences respectively; followed by the research data with 43, 45, and 59 times occurrences in Scopus, WoS and Merged file respectively.

Table-5: Top 10 Most Relevant Keywords

Scopus		WoS		Merged	
Keywords	Frequency	Keywords	Frequency	Keywords	Frequency
research data management	265	research data management	148	research data management	289
research data	43	research data	45	research data	59
data sharing	37	research	30	data management	43
data management	36	data	29	data sharing	40
open science	35	data management	29	open science	38

Lecture Notes in Computer Science including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics	28	Bibliothek Forschung Und Praxis	15	Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)	28
Communications in Computer & Information Science	21	IFLA Journal- International Federation of Library Associations	13	Communications in Computer and Information Science	21
CEUR Workshop Proceedings	19	DESIDOC Journal of Library and Information Technology	7	CEUR Workshop Proceedings	19
VOEB Mitteilungen	19	Program- Electronic Library and Information Systems	7	VOEB- Mitteilungen	19
Data Science Journal	15	Information Literacy in the Workplace	6	Bibliothek Forschung Und Praxis	15
ACM International Conf. Proceeding Series	12	Insights-The UKSG Journal	6	Data Science Journal	15
IFLA Journal	12	ISPRS International Journal of Geo- Information	6	IFLA Journal- International Federation of Library Associations	13
Liber Quarterly	12	Journal of Academic Librarianship	6	ACM International Conference Proceeding Series	12
Procedia Computer Science	12	Journal of The Medical Library Association	6	Liber Quarterly	12

Proceedings of the Association for Information Science and Technology	10	Journal of Documentation	5	Procedia Computer Science	12
---	----	--------------------------	---	---------------------------	----

The Table-7 depicts the top 10 most relevant affiliations on the basis of first author's affiliation for Scopus and WoS databases. It reveals that in Scopus, 51 publications (9.07%) were published in affiliation with University of Porto followed by University of Chicago with 29 publications (5.51%). But, in WoS 38 publications (10.33%) were published in affiliation with University of Porto followed by University of Sheffield with 24 publications (6.52%).

Table-7: Top 10 Most Relevant Affiliations

Scopus		WoS	
Affiliations	Documents	Affiliations	Documents
Universidade Do Porto (University of Porto)	51	Universidade Do Porto (University of Porto)	38
University of Chicago	29	University of Sheffield	24
University of Cologne	27	University of Cologne	15
University of Sheffield	26	University of Toronto	15
Universiti Putra Malaysia	24	University of Washington	12
Delft University of Technology	22	Delft University of Technology	11
University Medical Center GA-Ttingen (Georg August University)	22	Northumbria University	11
Universitat WIEN (University of Vienna)	19	Karlsruhe Institute of Technology	10
Bielefeld University	18	Dresden University of Technology	10
Imperial College London	18	University of Pittsburgh	10

5.6. Country Scientific Production

Analysis of the country's scientific production for both Scopus and WoS databases was based on the author's affiliated country. The Table-8 depicts the list of top 20 countries' contribution in globally published literatures on research data management during 2000 – 2019. By comparison, Table-8 reveals that the USA, Germany and the UK, these three countries were most productive countries in both Scopus and WoS databases, but afterwards there is inconsistency of countries in ranking in both the databases. The USA is having 445 published literatures in Scopus versus 294 published literatures in WoS. Germany and UK are having 382 and 173 published literatures in Scopus versus 224 and 118 published literatures in WoS. Portugal is on 4th position with 99 published literatures in Scopus whereas it is on 7th position with 43 published literatures in WoS.

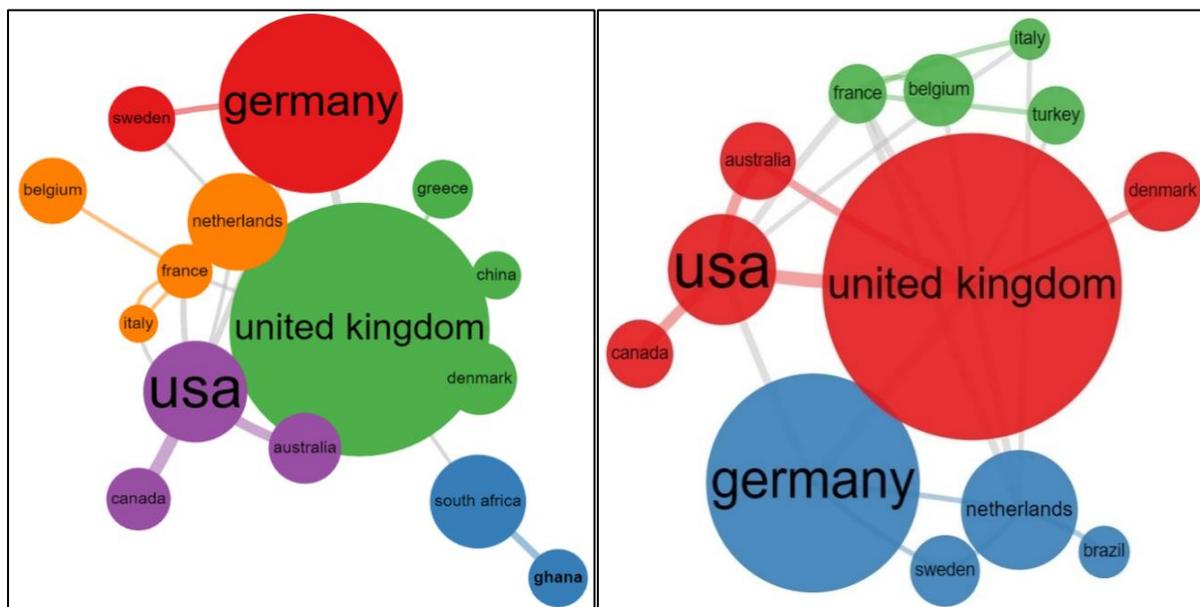
Table-8: Top 20 Countries on Scientific Production

Scopus		WoS	
Country	Documents	Country	Documents
USA	445	USA	294
Germany	382	Germany	224
UK	173	UK	118

Portugal	99	Canada	53
Australia	70	Australia	52
Austria	68	Netherlands	46
Netherlands	65	Portugal	43
Canada	56	China	32
China	49	India	31
France	34	South Africa	29
Malaysia	34	Switzerland	26
Sweden	29	Brazil	24
Switzerland	28	Japan	21
Japan	26	Spain	21
India	25	France	19
Belgium	22	Sweden	18
South Africa	17	Hungary	16
Brazil	15	Belgium	11
Italy	15	New Zealand	10
Spain	15	Turkey	10

5.7. International Collaboration

The Graph-11 and 12 depicts the collaboration between countries for both Scopus and WoS databases. There are 5 clusters in Scopus and 3 clusters in WoS data. The UK is having the largest node in both the databases, followed by Germany, thus the researchers from the UK and Germany were the top collaborators respectively with other countries in both the databases.



Graph-11 and 12: Countries Collaboration in Scopus and WoS

6. Conclusions

From the analysis it can be argued that Scopus covers research data management literature better than WoS. However, this study revealed that around 32% of globally published literatures on research data management were indexed in both the Scopus and WoS databases. Scopus contains 292 (55.5%) articles, followed by 194 Conference Paper/Conference Proceedings, whereas WoS contains 206 (55.98%) articles, followed by 132 (35.87) Conference Paper/Conference Proceedings. In the case of the annual percentage publications growth rates (Annual Growth rate), in Scopus it was 37.67%, whereas it was 39.67% in WoS which is higher than Scopus. The author named Ribeiro C was on 1st rank with 21 number of publications (NP) with an h-index of 6 in Scopus whereas he was on 2nd rank in WoS with 13 NP with an h-index of 3. The author named Da Silva JR was on 2nd rank with 17 NP with an h-index of 6 in Scopus whereas he was on 1st rank in WoS with 15 NP with an h-index of 3. The author named Castro JA was on 3rd rank with 16 NP in Scopus whereas he was on 4th rank in WoS with 8 NP with an h-index of 6 and 3 respectively. The author named Cox AM was having top rank in terms of h-index in both the databases with 8 NP in each database. It was revealed that in Scopus, 438 publications (83.27%) were multi-authored; whereas 303 publications (82.34%) were multi-authored. The most frequent keyword used was “research data management” which was on top in the list in both the databases, i.e. in Scopus with 265 occurrences and in WoS with 148 occurrences. The source “Lecture Notes in Computer Science including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics” was the top publication source in Scopus and the source “Bibliothek Forschung Und Praxis” was the top publication source in WoS. Further, the study presented that in Scopus, 51 literatures (9.07%) were published in affiliation with University of Porto followed by University of Chicago with 29 publications (5.51%); whereas in WoS, 38 publications (10.33%) were published in affiliation with University of Porto followed by University of Sheffield with 24 publications (6.52%). It revealed that the USA, Germany and the UK were the most productive countries in both Scopus and WOS databases; and the researchers from the UK and Germany were the top collaborators respectively with other countries in both the databases.

References

- Archambault, É., Campbell, D., & Larivière, Y. G. and V. (2013). Comparing Bibliometric Statistics Obtained From the Web of Science and Scopus. *Journal of the American Society for Information Science and Technology*, 64(July), 1852–1863. <https://doi.org/10.1002/asi>
- Bakkalbasi, N., Bauer, K., Glover, J., & Wang, L. (2006). Three options for citation tracking: Google Scholar, Scopus and Web of Science. *Biomedical Digital Libraries*, 3(1), 7. <https://doi.org/10.1186/1742-5581-3-7>
- Chirici, G. (2012). Assessing the scientific productivity of Italian forest researchers using the Web of Science, SCOPUS and SCIMAGO databases. *IForest - Biogeosciences and Forestry*, 5(1), 101–107. <https://doi.org/10.3832/ifor0613-005>
- Echchakoui, S. (2020). Why and how to merge Scopus and Web of Science during bibliometric analysis: the case of sales force literature from 1912 to 2019. *Journal of Marketing Analytics*, 8(3), 165–184. <https://doi.org/10.1057/s41270-020-00081-9>

Fernández, María Isabel Escalona; Barbosa, Pilar Lagar; Guerrero, A. P. (2010). Web of Science Vs. Scopus: Un Estudio Cuantitativo En Ingeniería Química. *Anales de Documentación*, 13(0), 159–175. <https://doi.org/10.6018/analesdoc.13.0.107121>

Gavel, Y., & Iselid, L. (2008). Web of Science and Scopus: a journal title overlap study. *Online Information Review*, 32(1), 8–21. <https://doi.org/10.1108/14684520810865958>

Mongeon, P., & Paul-Hus, A. (2016). The journal coverage of Web of Science and Scopus: a comparative analysis. *Scientometrics*, 106(1), 213–228. <https://doi.org/10.1007/s11192-015-1765-5>

Pritchard, A. (1969). Statistical Bibliography or Bibliometrics? *Journal of Documentation*, 25(4), 348–349.

Sánchez, A. D., de la Cruz Del Río Rama, M., & García, J. Á. (2017). Bibliometric analysis of publications on wine tourism in the databases Scopus and WoS. *European Research on Management and Business Economics*, 23(1), 8–15. <https://doi.org/10.1016/j.iedeen.2016.02.001>

Wagner, A. Ben. (2015). A Practical Comparison of Scopus and Web of Science Core Collection. Retrieved from <https://ubir.buffalo.edu/xmlui/handle/10477/38568>