TRENDS IN OPEN SCIENCE: A BIBLIOMETRIC ANALYSIS OF RESEARCH TOPICS, CITATIONS, JOURNALS, AND PRODUCTIVE ENTITIES

MAHFOOZ AHMED[1], ROSLINA BT. OTHMAN[1*], MOHAMAD FAUZAN BIN NOORDIN[2]

[1]*Kulliyyah of Information and Communication Technology, International Islamic University, Malaysia*

**Corresponding author: almahfooz4real@gmail.com*

ABSTRACT: A transformative shift has redefined how scientific research is conducted and shared in the digital era. Open Science is emerging as a transparent, collaborative, and accessible paradigm at the forefront of this revolution. This study undertakes a comprehensive Bibliometric analysis to explore the multifaceted landscape of Open Science, addressing key aspects such as prevalent topics, top-cited papers, leading journals, productive countries, and active institutions. Utilising a robust dataset from the Scopus online database, which encompasses scholarly literature, scientific publications, and open-access journals across diverse disciplines, the research employs advanced quantitative techniques to delineate the contours of Open Science. The findings reveal critical areas and topics, influential authors and works, prominent journals, and pioneering countries and institutions that shape the discourse. A geographical analysis further emphasises the importance of both national and international collaboration in the Open Science movement. Beyond contributing to academic understanding, this research provides essential insights for a wide range of stakeholders, from researchers to policymakers, and reflects on the broader implications of the findings. The paper concludes by projecting a path for future explorations, including emerging topics such as artificial intelligence, metadata, and ethics in Open Science, and offers lessons from the Open Science practices in response to the COVID-19 pandemic, making it a valuable guide for the ongoing evolution of scientific research and collaboration.

KEY WORDS: *Open Science; Bibliometric Analysis; Scopus database; Citation Analysis; Journal Analysis; Geographical Distribution*

1. INTRODUCTION

The advent of the digital era has revolutionised various facets of human life, and scientific research is no exception. The traditional barriers to accessing and sharing scientific knowledge are gradually being dismantled, giving rise to the Open Science movement. Open Science, characterised by transparency, collaboration, and accessibility, seeks to make scientific research, data, and dissemination available to all levels of society (McAbee et al., 2018). Open Science encompasses various components, including open access to publications, open research data, open-source software, open collaboration, and open peer review. These elements collectively contribute to a more inclusive and democratic scientific process, fostering innovation and accelerating the pace of discovery (Tennant et al., 2020).

With the open science system, researchers can work more on each other's findings and build on each other's work towards proving more sophisticated research findings. This system of research will also allow us to build a collaborative effort across disciplines, allowing new explanations and insights to emerge. Examples of these benefits can be seen in the current COVID-19 pandemic. In January 2020, researchers began sharing the genetic code of the SARS-CoV-2 virus with colleagues around the world to fight the pandemic. Edward Holmes, a professor at the University of Sydney, won the 2021 Prime Minister's Prize for Science for his role in this very act after he worked with colleagues in China and Scotland to release the genetic code, catalysing work on a test and a vaccine for the virus (Cathy, 2021).

The objective of open science is to establish a practice in which individuals could collaborate to contribute to all stages of the research process, and after the procedure, the research data, lab notes, and other research processes would be made publicly available under terms that allow for reuse, redistribution, and research reproduction. In contrast, the FAIR practice is involved in guiding the principles of increasing the findability, accessibility, interoperability, and reusability of digital data (Go Fair, 2016). The concept of open science also includes open access, which simply refers to free and unrestricted online access to research materials, findings such as in journal articles, books, book chapters, and other related open data, such as facts or statistics collected for reference or analysis, which anyone can freely access, use, modify, and share for any purpose; open source, which is related to open software, a computer application or programme source code; open educational resources, which is aligned with the idea of open culture and free sharing of knowledge and educational resources; and citizen science, which offers to transform everyone into a scientist, promising to supply new knowledge, educate the public, and reconfigure science from a closed to open activity, in brief, "democratising" science, etc. (FOSTER, 2020). Due to the high investment in constructing and maintaining a physical research laboratory, several research and technology organisations (RTO) have also been adopting the open lab model, which incorporates open design (OD) ideas with open data that address software and hardware design creation, using open-source software that is freely and digitally available and supported by communities of individuals with shared interests (Castro et al., 2021).

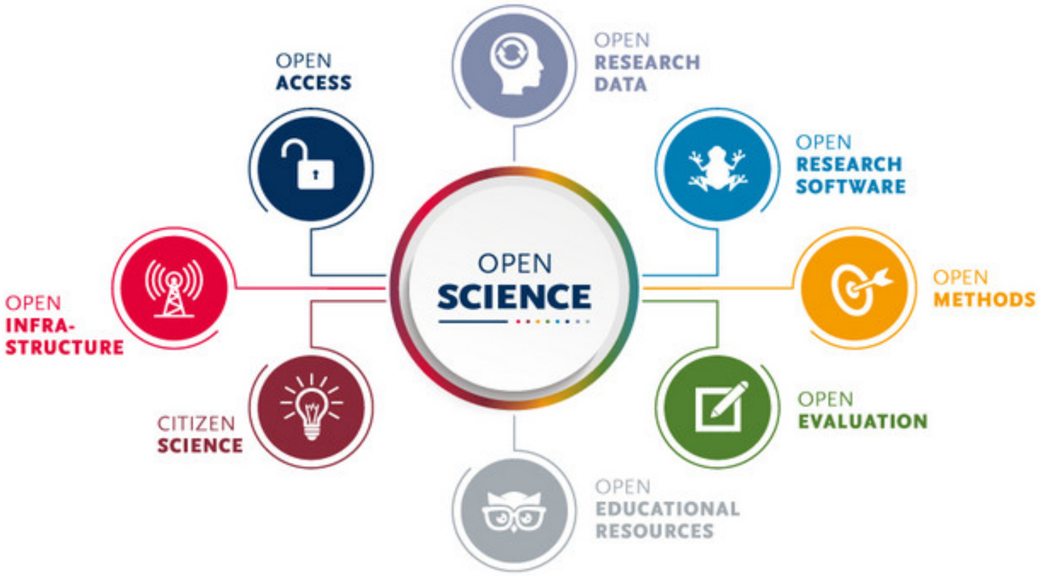


Fig. 1. Open science strategies; image source: the University of Potsdam.

Bibliometric analysis is a methodological approach that mostly applies quantitative analysis to scientific publications and offers a powerful tool to explore the dynamics of Open Science. By analysing patterns, trends, and networks within the scientific literature, bibliometric analysis provides insights into the most influential topics, journals, authors, and regions in a specific field of study (Aria & Cuccurullo, 2017). This paper aims to search into the evolving landscape of Open Science through a comprehensive bibliometric analysis. Utilising a dataset from the Scopus online database, which is a robust information source encompassing scholarly literature across various disciplines, the study seeks to answer five fundamental questions related to the prevalence of topics, citations, leading journals, productive countries, and active institutions in Open Science research specifically written as follows.

1. What are the most popular topics in open science research?
2. What are the most cited open science papers?
3. What are the most influential journals in open science research?
4. What are the most productive countries in open science research?
5. What are the most productive institutions in open science research?

The following sections will further elaborate on the methodology used in collecting the applicable data of the study, analysing the data, presenting the findings, and discussing the implications of this study, contributing to the broader understanding of Open Science and its role in shaping the future of scientific research).

2. METHODOLOGY

2.1. Data Collection and Selection Criteria

The primary dataset for this study was sourced from the Scopus online database, a comprehensive and robust information source encompassing scholarly literature across various disciplines (Burnham, 2006). The database was queried for publications related to Open Science using a single query of the phrase “open science,” where the phrase appeared in the articles’ titles, abstracts, or keywords. The search was refined to include only studies conducted from 2013 to 2023, encompassing all document types, such as journal articles, conference papers, books, and book chapters. Both final published articles and articles in the press were included, covering all journals currently indexed in the database. The search also encompassed authors’ affiliations, funding sponsors, countries/territories, source types, and languages. The Scopus search result interface is provided in Figure 2 below.

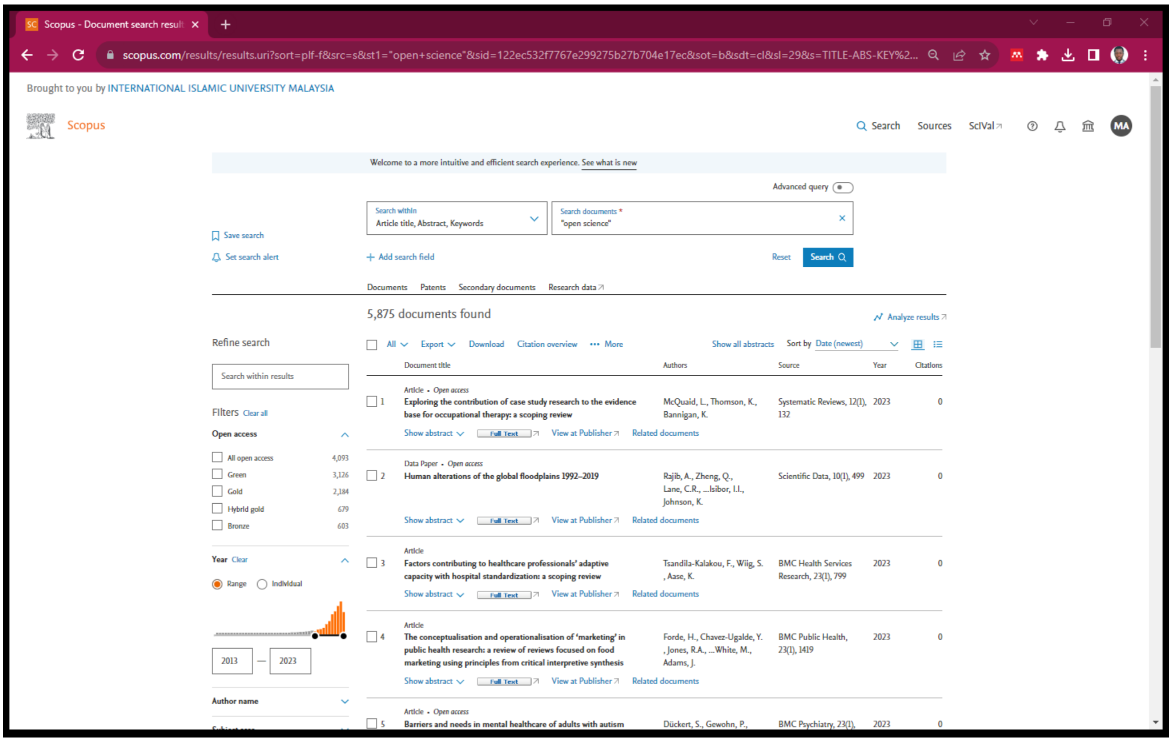


Fig. 2. Scopus search results interface.

2.2. Bibliometric Analysis

Bibliometric analysis is a quantitative approach to the study of scientific and technological literature. It involves the statistical analysis of books, articles, and other related publications, with the aim of understanding the patterns, connections, and trends within a specific field or subject (Donthu et al., 2021). Bibliometric methods are predominantly used to analyze the impact of a field, the impact of a set of researchers, or the impact of a particular paper, by examining aspects such as citation counts, authorship, and publication outlets (Roldan-Valadez et al., 2019).

In the context of open science, bibliometric analysis can provide valuable insights into the global landscape of research collaboration, influential authors, and emerging trends. For example, a bibliometric study by MANGHI P and SCHÖPFEL J (as seen in the provided data) reveals the most cited authors and journals in the field, highlighting the key areas of focus and collaboration patterns. The fractionalized counts of citations can also shed light on individual and shared contributions to the field, reflecting different approaches to collaboration and research (Donthu et al., 2021).

However, bibliometric analysis is not without its challenges and limitations. In many instances, it requires careful consideration of the data sources, methods, and metrics used, as biases and disparities may affect the results. For instance, the applicable dataset used in this study was only retrieved from the Scopus database. Likewise, several other scientific databases may be out there with different results. Despite these challenges, bibliometric analysis remains a powerful tool for understanding the scientific landscape, guiding research policy, and informing decisions related to promotions, tenure, and funding. Its application to open science underscores the importance of transparency, accessibility, and collaboration in the pursuit of knowledge and innovation. This is why bibliometric analysis was employed to explore the dynamics of open science through the following quantitative steps:

- Topic Analysis: Identification of the most popular topics in Open Science research through keyword frequency analysis and topic modelling.
- Citation Analysis: Examination of the most cited Open Science papers, utilising citation counts and impact metrics.
- Journal Analysis: Analysis of the most influential journals in Open Science research, based on factors such as impact factor and journal ranking.
- Country and Institution Analysis: Geographical analysis to identify the most productive countries and institutions, using publication counts, collaboration networks, and regional contributions).

2.3.Tools and Software

The analysis was conducted using specialized bibliometric tools and software, including RStudio biblioshiny software, VOSviewer, and the Scopus basic statistical analysed search results. These tools were selected for distinct reasons. RStudio biblioshiny software is widely recognized for its flexibility and efficiency in handling extensive bibliometric datasets, offering a user-friendly interface for data manipulation and visualization, thus making it apt for intricate analyses (Aria & Cuccurullo, 2017). Similarly, VOSviewer has gained popularity as a tool for constructing and visualizing bibliometric networks, enabling the mapping of co-authorship, co-citation, and keyword co-occurrence networks. This provides a lucid view of collaboration patterns and thematic structures within a field (Perianes-Rodriguez et al., 2016). Lastly, since the applicable dataset used in the analysis was retrieved from the Scopus online database, it was deemed necessary to include some statistical results figures and tables directly from the database. These tools were chosen for their collective ability to facilitate data visualization, network analysis, and statistical modelling, allowing for a comprehensive exploration of the open science landscape. Their combined utilization leveraged the strengths of each tool, offering a multifaceted approach to understanding the dynamics of open science research.

Fig. 3. Tools and Software.

3. ANALYSIS AND RESULTS VISUALISATION

3.1. Topic Analysis

Topic analysis, also known as topic modelling, is a method primarily used to discover the abstract “topics” that occur in a collection of documents (Albalawi et al., 2020). It’s a form of text mining that has become increasingly essential in various fields, from academic research to business intelligence. In academic research, for example, topic analysis is often employed to uncover underlying themes in large datasets of textual information, revealing trends and patterns in specific research domains (Fischer et al., 2020). In the business world, topic analysis can be used to understand customer feedback, social media conversations, and product reviews (Ibrahim & Wang, 2019). By analysing recurring themes in customer comments, companies can gain insights into product strengths and weaknesses. Online shopping centres, for instance, might utilise topic modelling to categorise product reviews, helping both consumers and sellers understand the key features people discuss. Sectors like healthcare, journalism, and media can also use topic analysis to analyse medical records and research articles, identify trends in patient care, make data-driven decisions, or track the evolution of news stories and public opinion.

To conduct the topic analysis of the most frequent themes within the open science research, the RStudio biblioshiny software was utilised on the bibliographic dataset obtained from the Scopus database. This analysis uncovered several key focus areas in open science research, with the top fifty topics presented in the document analysis TreeMap and the accompanying table below.

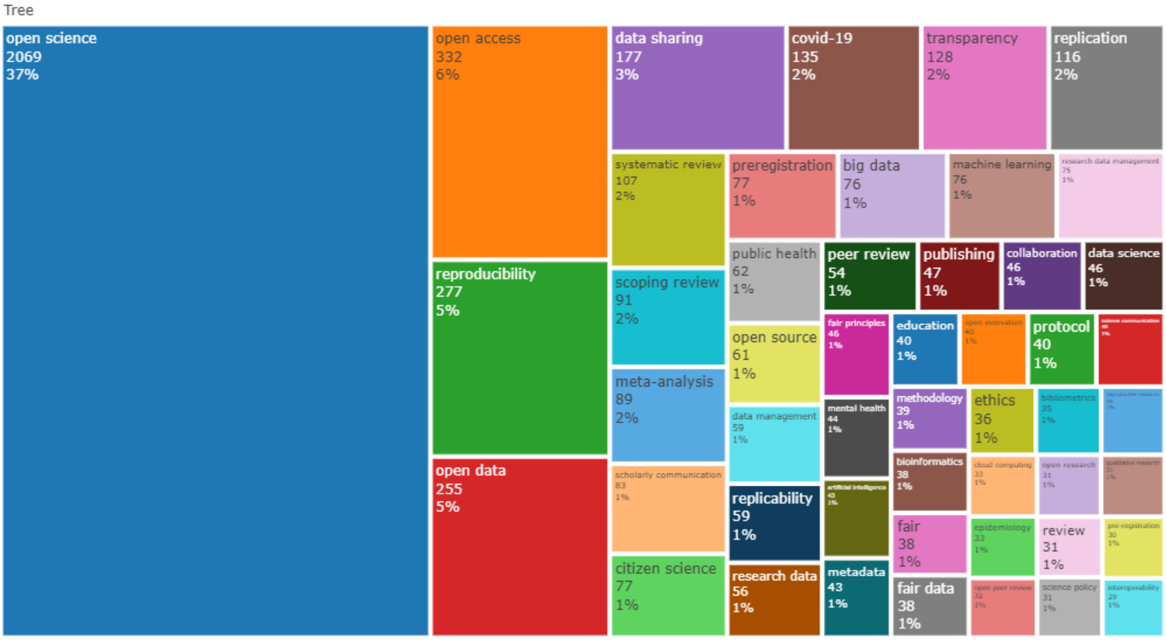


Fig. 4. Topic analysis TreeMap.

Table 1: Topic analysis table

S/N	Words	Occu.	S/N	Words	Occu.
1	open science	2069	26	data science	46
2	open access	332	27	fair principles	46
3	reproducibility	277	28	mental health	44
4	open data	255	29	artificial intelligence	43
5	data sharing	177	30	metadata	43
6	covid-19	135	31	education	40
7	transparency	128	32	open innovation	40
8	replication	116	33	protocol	40
9	systematic review	107	34	science communication	40
10	scoping review	91	35	methodology	39
11	meta-analysis	89	36	bioinformatics	38
12	scholarly communication	83	37	fair data	38
13	citizen science	77	38	ethics	36
14	preregistration	77	39	bibliometrics	35
15	big data	76	40	reproducible research	34
16	machine learning	76	41	cloud computing	33
17	research data management	75	42	epidemiology	33
18	public health	62	43	open peer review	32
19	open source	61	44	open research	31
20	data management	59	45	qualitative research	31
21	replicability	59	46	science policy	31
22	research data	56	47	pre-registration	30
23	peer review	54	48	interoperability	29
24	publishing	47	49	publication bias	29
25	collaboration	46	50	research methods	29

According to the topic analysis results in the table, the most frequent occurrence words after "open science" are "open access," "reproducibility," "open data," and "data sharing." These are undoubtedly the most trending and focused areas (topics) in open science at present. Open science represents a structural research approach that promotes practices of openness, integrity, and reproducibility (Banks et al., 2019). Within the open science framework, researchers can build on each other's findings, fostering collaboration across disciplines and allowing new insights to emerge. The following paragraphs provide a brief explanation of why these topics are becoming more prevalent:

- Open Access: The open access system is known for accelerating knowledge transmission by making published content freely shared (Gallagher et al., 2020). It allows free and unrestricted internet access to research findings, such as journal articles, books, and book chapters. Typically, open-access content can be viewed without charge, referring to digital, peer-reviewed scientific publications that are free to read, with minimal copyright and licensing limitations (Suber, 2010). To meet the open science objectives of promoting openness, integrity, and reproducibility in research and knowledge sharing, open access practice needs to be promoted. This could explain why numerous studies are being conducted on this topic to advance open access initiatives in various publication and knowledge areas.

- Reproducibility: Reproducibility in research refers to the act of making original research data, analysis code, and other materials available as part of the journal submission. The availability of original data supports the reproducibility of reported research results, as other investigators can analyse the data independently (Spitschan et al., 2021). Research is considered reproducible when others can reproduce the results using only the original data, code, and documentation (Essawy et al., 2020). In line with the primary objective of the open science system to promote openness, integrity, and reproducibility in research and knowledge sharing (Banks et al., 2019), researchers are increasingly focusing on making original research data reproducible to further promote scientific and knowledge integrity.

- Open Data: Similarly, open data or data sharing refers to the process of making original research data available to others, either upon completing the research or upon request (Ahmed & Othman, 2021). Scientific research in the 21st century has become more data-intensive and collaborative, making it essential to study data practices related to preservation, discovery, accessibility, and reuse. Data sharing and reuse are valuable parts of the scientific method, allowing for verification of results and extending research from prior results (Tenopir et al., 2020).

- COVID-19: The COVID-19 pandemic, caused by the SARS-CoV-2 virus, emerged in Wuhan, China, in late 2019 and quickly escalated into a global crisis, affecting millions of lives and healthcare systems worldwide (Cardenas-Gonzalez & Alvarez-Buylla, 2020). The rapid spread led to unprecedented public health measures, including widespread lockdowns and travel restrictions. An example of open science benefits during the pandemic includes researchers sharing the genetic code of the SARS-CoV-2 virus to fight the pandemic, with Edward Holmes winning the 2021 Prime Minister’s Prize for Science for his role in releasing the genetic code, catalysing work on a test and vaccine for the virus (Cathy, 2021). Since then, many researchers have come to recognize the numerous benefits of the open science research system, encouraging global collaboration among researchers in support of transparency and reproducible research to further save costs.

These topics reflect the current trends in open science, emphasizing the importance of collaboration, transparency, and accessibility in the pursuit of knowledge and innovation).

3.2. Citation Analysis

Citation analysis examines the frequency, patterns, and graphs of document citations. It provides insights into the relationships between authors, works, and disciplines and is vital in bibliometrics and scientometrics (Tomaszewski, 2023). Citation analysis plays a crucial role in academic research by measuring the impact and relevance of scholarly works. By analysing how often a paper is cited, researchers can gauge its influence within a particular field. The h-index, introduced by Hirsch, (2005), is a popular metric that quantifies both the productivity and impact of a scientist's publications. Citation analysis can be used for several purposes. For example, libraries can use it to build their collections by understanding which journals are most frequently cited, ensuring that they subscribe to the most relevant and influential journals in a particular field (Enger, 2009). Scientific research funding agencies and policymakers can use citation analysis to identify core and emerging trends in specific research areas and assess the research's impact on society.

Examining the most cited open science papers provides insight into the influential works shaping the field. Identifying the most influential authors, documents, and trends in this study area will greatly help shape future research directions. This analysis will evaluate the researchers and their research performance, helping in understanding their academic influence while informing decisions related to promotions, tenure, and funding. Figure 4 and Table 2 below list the top-cited authors and documents, showcasing the impact of their works on the entire research lifecycle.

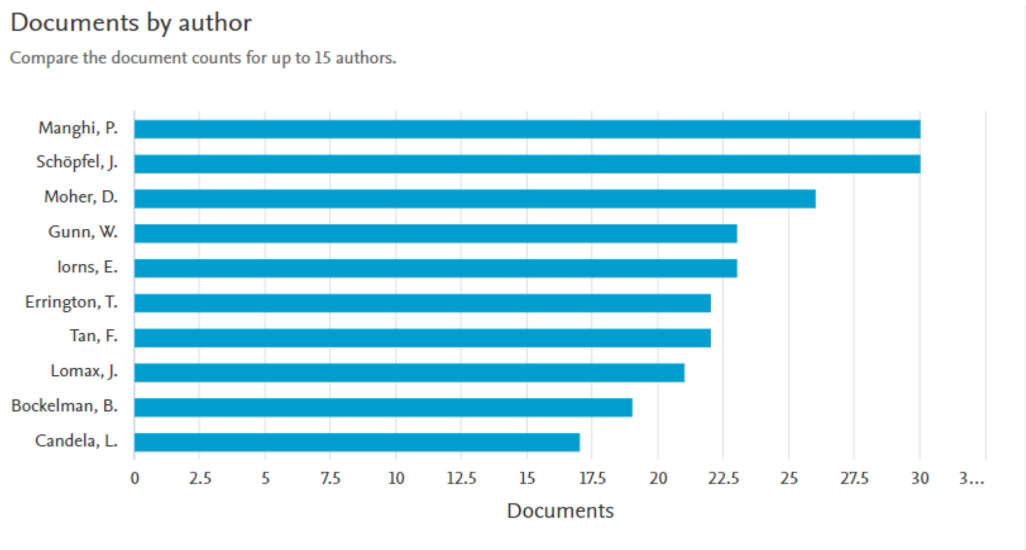


Fig. 5. The top 10 cited authors

Table 2: The top 50 cited authors

S/N	Authors	Articles	Articles Fractio-nalized	S/N	Authors	Articles	Articles Fractio-nalized
1	MANGHI P	30	6.56	26	HUANG X	14	1.98
2	SCHÖPFEL J	30	13.58	27	KIM J	14	3.26
3	CHEN Y	28	3.81	28	ZHANG H	14	1.18
4	MOHER D	26	4.75	29	ZHANG Y	14	1.98
5	GUNN W	23	2.55	30	HUANG J	13	2.92
6	IORNS E	23	2.55	31	LI L	13	2.34
7	WANG Y	23	3.66	32	LIU Y	13	2.07
8	ERRINGTON T	22	2.38	33	BARDI A	12	3.59
9	TAN F	22	2.38	34	CHEN J	12	1.52
10	ZHANG L	22	3.64	35	DAS S	12	1.04
11	LOMAX J	21	2.28	36	IOANNIDIS JPA	12	1.96
12	BOCKELMAN B	19	3.34	37	LIU J	12	1.41
13	LI J	19	2.55	38	ROSS JS	12	4.03
14	WANG X	19	5.47	39	WANG H	12	1.59
15	ZHANG X	18	2.82	40	LI H	11	1.78
16	CANDELA L	17	3.11	41	LI Y	11	1.63
17	LIU X	17	2.68	42	LIU H	11	1.34
18	WANG L	17	2.17	43	MAYO-WILSON E	11	1.76
19	ZHANG J	17	2.33	44	NILSONNE G	11	0.90
20	CHEN X	16	2.02	45	ROSS-HELLAUER T	11	3.64
21	CORO G	16	5.18	46	ZHANG Z	11	2.25
22	PAGANO P	16	2.96	47	CHEN Z	10	1.95
23	PROST H	16	5.67	48	COOK BG	10	2.52
24	WANG J	16	2.15	49	KRUMHOLZ HM	10	1.88
25	PERFITO N	15	1.51	50	NAUDET F	10	2.40

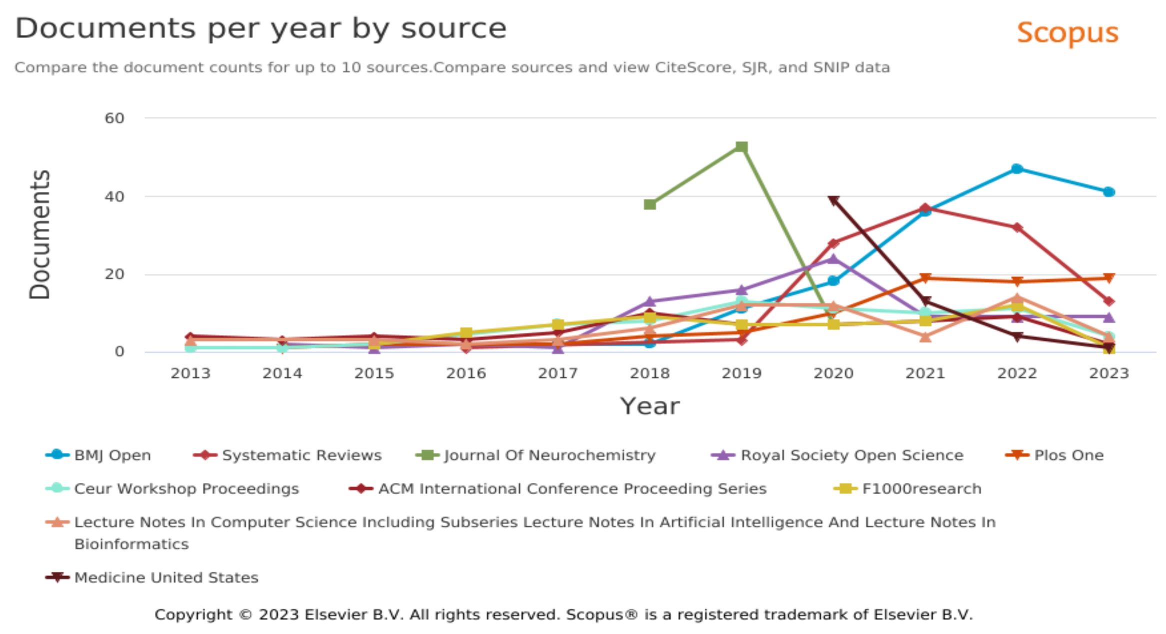
Both Figure 4 and Table 2 above present a snapshot of the top-cited authors in the open science research domain, highlighting the number of articles cited and a fractionalised count representing individual and collaborative contributions (Sivertsen et al., 2019). Fractionalised counts provide a nuanced view of an author's contributions, taking into account shared authorship in collaborative works. For instance, authors MANGHI P and SCHÖPFEL J lead the citation rank with 30 cited articles each, though with varying fractionalised counts of 6.56 and 13.58, respectively, reflecting different collaboration patterns. The rest of the authors listed in the table, from 1 to 50, have also been cited in between 10 to 28 articles, with fractionalised counts ranging from 1.51 to 5.18. This dataset provides insights into these authors' influential and collaboration dynamics, revealing a mix of individual and shared contributions to the field. The variations in the fractionalised counts across the authors may indicate differing approaches to collaboration and individual research within this study area.

3.3. Journal Analysis

The journal analysis is another critical aspect of the bibliometrics and scientometrics, focusing on evaluating and assessing academic journals (Tomaszewski, 2023). It encompasses various metrics and methods to understand journals' quality, impact, and influence within specific disciplines. The journal analysis is essential for researchers, librarians, publishers, and policymakers, providing insights into the scholarly landscape (Agarwal et al., 2016). Journal analysis can help determine a particular journal's quality and reputation in several instances. Researchers can identify the most prestigious journals in their field by evaluating factors such as the impact factor, citation rates, and peer-review process (Owan & Owan, 2021).

Garfield (1972) introduced the journal impact factor to measure the average number of citations received by articles in a journal during a specific period. Impact factor has today become a widely used metric for assessing journal influence (Garfield, 1994). Apart from the impact factor, a few other metrics can be used while considering the significance of a particular journal, such as the eigenfactor metrics to consider the quality of the citing journals, providing a more nuanced understanding of a journal's influence within the scholarly network, and the altimetric to capture online attention and engagement, including social media mentions, blog posts, and news coverage (Kunze et al., 2020).

Several benefits are also attached to journal analysis: to analyse the growth and impact of open-access journals to reveal the trends in publication fees, subject areas, and citation rates and to understand interdisciplinary Journals. By employing a combination of traditional metrics like impact factors and innovative approaches like altmetrics, journal analysis will provide a comprehensive view of the academic landscape, guide researchers in selecting appropriate publication venues, assist librarians in curating collections, and helps publishers in enhancing their journals' profiles. As the scholarly communication ecosystem continues to evolve, journal analysis remains a vital tool for understanding and navigating the complex world of academic publishing. The analysis of the influential journals in open science research here also identifies several leading publications, such as the top journals, their impact factors, rankings, and contributions to the open science research area.



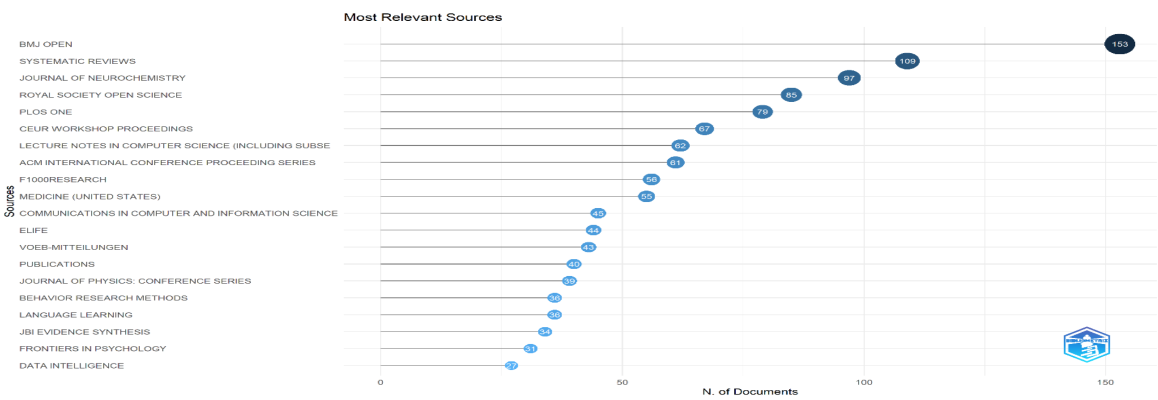


Fig. 6. Most Relevant Sources

Table 3: Most Relevant Sources

S/N	Sources	Articles	S/N	Sources	Articles
1	BMJ OPEN	153	26	SCIENTIFIC DATA	25
2	SYSTEMATIC REVIEWS	109	27	SCIENTOMETRICS	24
3	JOURNAL OF NEUROCHEMISTRY	97	28	PEERJ	23
4	ROYAL SOCIETY OPEN SCIENCE	85	29	ADVANCES IN METHODS AND PRACTICES IN PSYCHOLOGICAL SCIENCE	21
5	PLOS ONE	79	30	ECOLOGY AND EVOLUTION	21
6	CEUR WORKSHOP PROCEEDINGS	67	31	NEUROIMAGE	21
7	LECTURE NOTES IN COMPUTER SCIENCE (INCLUDING SUBSERIES LECTURE NOTES IN	62	32	PROCEEDINGS OF SCIENCE	21

	ARTIFICIAL INTELLIGENCE AND LECTURE NOTES IN BIOINFORMATICS)				
8	ACM INTERNATIONAL CONFERENCE PROCEEDING SERIES	61	33	GIGASCIENCE	20
9	F1000RESEARCH	56	34	CIENCIA DA INFORMACAO	19
10	MEDICINE (UNITED STATES)	55	35	COLLABRA: PSYCHOLOGY	19
11	COMMUNICATIONS IN COMPUTER AND INFORMATION SCIENCE	45	36	JOURNAL OF SYSTEMS AND SOFTWARE	19
12	ELIFE	44	37	CONFERENCE ON HUMAN FACTORS IN COMPUTING SYSTEMS - PROCEEDINGS	17
13	VOEB-MITTEILUNGEN	43	38	GREY JOURNAL	17
14	PUBLICATIONS	40	39	NATURE	16
15	JOURNAL OF PHYSICS: CONFERENCE SERIES	39	40	PLOS BIOLOGY	16
16	BEHAVIOR RESEARCH METHODS	36	41	PROCEEDINGS OF THE ASSOCIATION FOR INFORMATION SCIENCE AND TECHNOLOGY	16
17	LANGUAGE LEARNING	36	42	SPORTS MEDICINE	16
18	JBI EVIDENCE SYNTHESIS	34	43	PERSPECTIVES ON PSYCHOLOGICAL SCIENCE	15
19	FRONTIERS IN PSYCHOLOGY	31	44	INTERNATIONAL ARCHIVES OF THE PHOTOGRAMMETRY, REMOTE SENSING AND SPATIAL INFORMATION SCIENCES - ISPRS ARCHIVES	14
20	DATA INTELLIGENCE	27	45	JOURNAL OF LIBRARY AND INFORMATION SCIENCE IN AGRICULTURE	14
21	PROCEDIA COMPUTER SCIENCE	27	46	DIGITAL PRESENTATION AND PRESERVATION OF CULTURAL AND SCIENTIFIC HERITAGE	13
22	INFORMATION SERVICES AND USE	26	47	FRONTIERS IN NEUROINFORMATICS	13
23	DATA SCIENCE JOURNAL	25	48	FRONTIERS IN NEUROSCIENCE	13
24	GL-CONFERENCE SERIES: CONFERENCE PROCEEDINGS	25	49	GEOSCIENCE DATA JOURNAL	13
25	LIBER QUARTERLY	25	50	JOURNAL OF CLINICAL EPIDEMIOLOGY	13

Figure 5 and Table 3 above showcase the top 50 sources or journals that have published articles related to open science, ranging across various disciplines. Leading the list are “BMJ OPEN” with 153 articles and “SYSTEMATIC REVIEWS” with 109 articles, reflecting the strong presence of medical and healthcare research in open science. Other notable inclusions are multidisciplinary journals like “ROYAL SOCIETY OPEN SCIENCE” and “PLOS ONE,” computer science series such as “LECTURE NOTES IN COMPUTER SCIENCE,” and innovative platforms like “F1000RESEARCH.” The list also includes specialised journals in fields like neurochemistry, psychology, data science, and epidemiology. The diversity of these sources highlights the interdisciplinary nature of open science and its growing influence across different research fields. The presence of both traditional high-impact journals and innovative open-access platforms indicates a broad acceptance of open science principles and a shift towards more transparent and collaborative scientific research.

Table 4: Sources’ Local Impact

S/N	Element	h_index	g_index	m_index	TC	NP	PY_start
1	JOURNAL OF NEUROCHEMISTRY	21	28	3.5	1502	97	2018
2	LANGUAGE LEARNING	15	27	2.5	750	36	2018
3	BMJ OPEN	14	20	2	546	153	2017
4	ADVANCES IN METHODS AND PRACTICES IN PSYCHOLOGICAL SCIENCE	13	21	2.167	667	21	2018
5	BEHAVIOR RESEARCH METHODS	13	36	1.625	2118	36	2016
6	F1000RESEARCH	13	28	1.444	814	56	2015
7	PEERJ	13	23	1.444	979	23	2015
8	ROYAL SOCIETY OPEN SCIENCE	13	23	1.3	571	85	2014
9	ELIFE	12	34	1.2	1183	44	2014
10	FRONTIERS IN PSYCHOLOGY	12	16	1.2	282	31	2014
11	NEUROIMAGE	11	21	1	1047	21	2013
12	PLOS ONE	11	27	1.222	764	79	2015
13	PUBLICATIONS	10	16	1.25	310	40	2016
14	RESEARCH POLICY	10	12	0.909	372	12	2013
15	GIGASCIENCE	9	20	1	476	20	2015
16	PERSPECTIVES ON PSYCHOLOGICAL SCIENCE	9	15	0.818	827	15	2013
17	PLOS BIOLOGY	9	16	0.9	592	16	2014
18	ECOLOGY AND EVOLUTION	8	12	0.889	175	21	2015
19	FRONTIERS IN NEUROINFORMATICS	8	13	0.8	328	13	2014
20	FRONTIERS IN NEUROSCIENCE	8	13	0.727	185	13	2013
21	LECTURE NOTES IN COMPUTER SCIENCE (INCLUDING SUBSERIES LECTURE NOTES IN	8	17	0.727	349	62	2013

	ARTIFICIAL INTELLIGENCE AND LECTURE NOTES IN BIOINFORMATICS)						
22	NEURON	8	8	1	263	8	2016
23	SCIENTIFIC DATA	8	25	0.8	694	25	2014
24	SCIENTOMETRICS	8	17	0.727	300	24	2013
25	SYSTEMATIC REVIEWS	8	12	1	291	109	2016

Table 4 provides an insightful overview of the local impact of various sources in the field of open science, as measured by several bibliometric indicators. The “JOURNAL OF NEUROCHEMISTRY” leads with an h-index of 21, a g-index of 28, and an m-index of 3.5, reflecting a strong influence in the field since 2018. Other notable sources include “LANGUAGE LEARNING” and “BMJ OPEN,” both demonstrating significant impact through their respective h, g, and m indices. The table also highlights the presence of diverse journals, ranging from “ADVANCES IN METHODS AND PRACTICES IN PSYCHOLOGICAL SCIENCE” to “FRONTIERS IN NEUROSCIENCE” and “SCIENTOMETRICS.” These metrics collectively provide a snapshot of the influence and reach of these sources within the scientific community. The variation in the starting publication year (PY_start) among these sources, ranging from 2013 to 2018, further illustrates the evolving landscape of open science and the growing recognition of various journals in promoting transparent and collaborative research.

3.4. Country and Institution Analysis

Country and Institution Analysis is also vital to bibliometrics and research evaluation, providing insights into scientific research’s geographical and organizational distribution. This analysis helps understand the collaboration patterns, productivity, impact, and specialization of countries and institutions in various scientific domains (Kodonas et al., 2021). By analyzing the research output of countries and institutions, policymakers and researchers can easily identify the leading players in specific fields and assess their global standing. Analyzing the number of publications and citations helps assess the research productivity and impact of countries and institutions while mapping the collaboration networks between countries and institutions provides insights into the structure and dynamics of global research collaboration.

Country and Institution Analysis is a powerful tool for mapping the global landscape of scientific research. It provides valuable insights into the productivity, impact, collaboration, and specialization of countries and institutions. This analysis guides policymakers, funding agencies, and research institutions in making informed decisions, fostering collaboration, and enhancing the quality and relevance of research. As the global research ecosystem becomes increasingly interconnected and competitive, Country and Institution Analysis remains essential for understanding the dynamics of scientific innovation and excellence. Here, the geographical analysis highlighted the most productive countries and institutions in open science research.

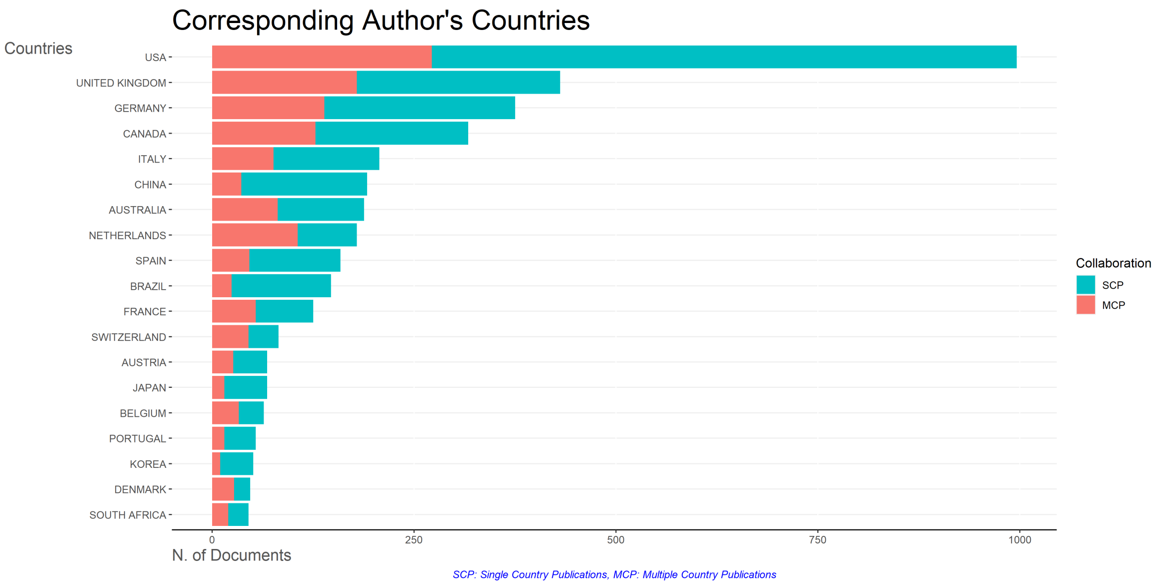


Fig. 7. Corresponding Author's Countries.

Table 5: Corresponding Author's Countries

S/N	Country	Articles	SCP	MCP	Freq	MCP_Ratio
1		1478	1156	322	0.254	0.218
2	USA	996	724	272	0.171	0.273
3	UNITED KINGDOM	431	252	179	0.074	0.415
4	GERMANY	375	236	139	0.064	0.371
5	CANADA	317	189	128	0.054	0.404
6	ITALY	207	131	76	0.036	0.367
7	CHINA	192	156	36	0.033	0.188
8	AUSTRALIA	188	107	81	0.032	0.431
9	NETHERLANDS	179	73	106	0.031	0.592
10	SPAIN	159	113	46	0.027	0.289
11	BRAZIL	147	123	24	0.025	0.163
12	FRANCE	125	71	54	0.021	0.432
13	SWITZERLAND	82	37	45	0.014	0.549
14	AUSTRIA	68	42	26	0.012	0.382
15	JAPAN	68	53	15	0.012	0.221
16	BELGIUM	64	31	33	0.011	0.516
17	PORTUGAL	54	39	15	0.009	0.278
18	KOREA	51	41	10	0.009	0.196
19	DENMARK	47	20	27	0.008	0.574
20	SOUTH AFRICA	45	25	20	0.008	0.444
21	SWEDEN	43	14	29	0.007	0.674
22	NORWAY	40	14	26	0.007	0.65
23	FINLAND	39	27	12	0.007	0.308
24	NEW ZEALAND	34	21	13	0.006	0.382
25	IRELAND	32	20	12	0.005	0.375
26	GREECE	26	16	10	0.004	0.385

27	POLAND	24	17	7	0.004	0.292
28	INDIA	23	15	8	0.004	0.348
29	MEXICO	23	15	8	0.004	0.348
30	SINGAPORE	16	9	7	0.003	0.438
31	GEORGIA	15	8	7	0.003	0.467
32	ARGENTINA	14	9	5	0.002	0.357
33	CZECH REPUBLIC	14	10	4	0.002	0.286
34	MALAYSIA	14	5	9	0.002	0.643
35	UKRAINE	14	12	2	0.002	0.143
36	COLOMBIA	13	7	6	0.002	0.462
37	CROATIA	13	9	4	0.002	0.308
38	CUBA	11	8	3	0.002	0.273
39	IRAN	9	5	4	0.002	0.444
40	TURKEY	8	5	3	0.001	0.375
41	HONG KONG	7	4	3	0.001	0.429
42	ISRAEL	7	1	6	0.001	0.857
43	SLOVENIA	7	6	1	0.001	0.143
44	LUXEMBOURG	6	1	5	0.001	0.833
45	NIGERIA	6	2	4	0.001	0.667
46	THAILAND	6	3	3	0.001	0.5
47	CHILE	4	2	2	0.001	0.5
48	ECUADOR	4	2	2	0.001	0.5
49	HUNGARY	4	3	1	0.001	0.25
50	INDONESIA	4	1	3	0.001	0.75

Table 5 offers a comprehensive view of the distribution of corresponding authors' countries in the field of open science research. The unnamed country at the top of the list leads with 1478 articles, followed by the USA with 996 articles, and the United Kingdom with 431. The table also highlights the Single Country Publications (SCP) and Multi-Country Publications (MCP) for each nation, along with the frequency (Freq) and MCP ratio. Countries like the Netherlands, Sweden, and Norway exhibit a higher MCP ratio, indicating a strong inclination towards international collaboration in research. Conversely, countries like Brazil, China, and Korea have a lower MCP ratio, reflecting a more localised approach to research collaboration. The table spans a diverse range of countries, from major research hubs to emerging contributors, providing a global perspective on the collaborative dynamics and research output in open science. This data underscores the importance of both national and international collaboration in shaping the landscape of open science research.

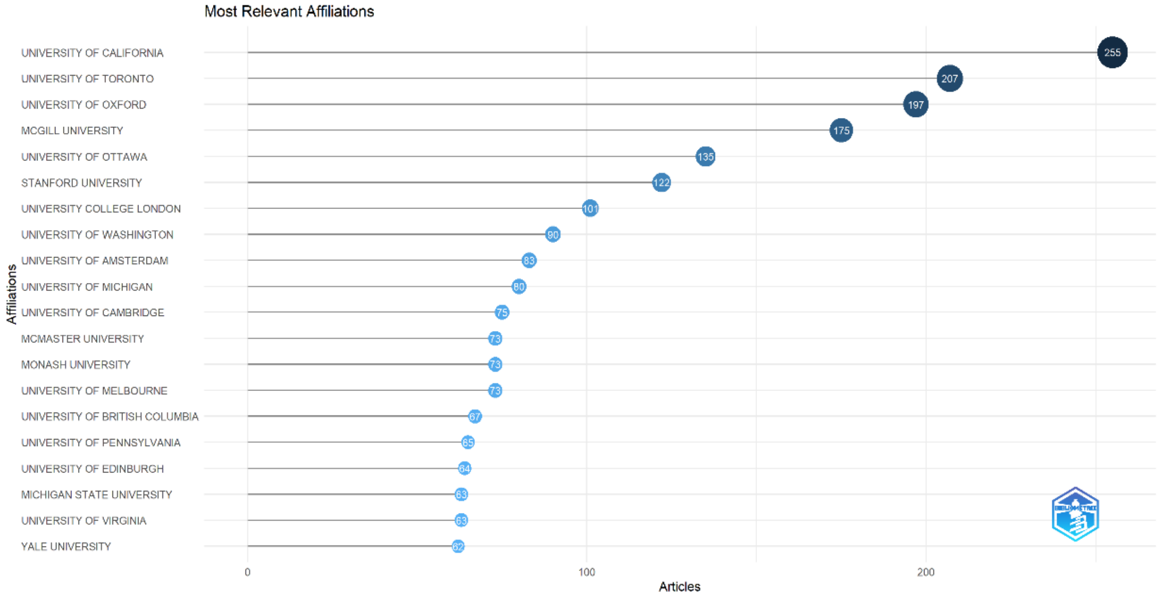


Fig. 8. Most Relevant Affiliations.

Table 6: Most Relevant Affiliations)

S/N	Affiliation	Articles	S/N	Affiliation	Articles
1	UNIVERSITY OF CALIFORNIA	255	26	UNIVERSITY OF CALGARY	57
2	UNIVERSITY OF TORONTO	207	27	COLUMBIA UNIVERSITY	56
3	UNIVERSITY OF OXFORD	197	28	UNIVERSITY OF SOUTHERN CALIFORNIA	56
4	MCGILL UNIVERSITY	175	29	ARIZONA STATE UNIVERSITY	54
5	UNIVERSITY OF OTTAWA	135	30	INDIANA UNIVERSITY	53
6	STANFORD UNIVERSITY	122	31	IMPERIAL COLLEGE LONDON	51
7	UNIVERSITY COLLEGE LONDON	101	32	UNIVERSITY OF BRISTOL	51
8	UNIVERSITY OF WASHINGTON	90	33	UNIVERSITY OF MANITOBA	51
9	UNIVERSITY OF AMSTERDAM	83	34	UNIVERSITY OF ALBERTA	50
10	UNIVERSITY OF MICHIGAN	80	35	DALHOUSIE UNIVERSITY	49
11	UNIVERSITY OF CAMBRIDGE	75	36	UNIVERSITY OF COPENHAGEN	49
12	MCMASTER UNIVERSITY	73	37	UNIVERSITY OF NEW SOUTH WALES	49
13	MONASH UNIVERSITY	73	38	DUKE UNIVERSITY	48
14	UNIVERSITY OF MELBOURNE	73	39	UNIVERSITY OF ARIZONA	48
15	UNIVERSITY OF BRITISH COLUMBIA	67	40	VRIJE UNIVERSITEIT AMSTERDAM	48
16	UNIVERSITY OF PENNSYLVANIA	65	41	CENTER FOR OPEN SCIENCE	47
17	UNIVERSITY OF EDINBURGH	64	42	UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN	47
18	MICHIGAN STATE UNIVERSITY	63	43	UNIVERSITY OF SYDNEY	46

19	UNIVERSITY OF VIRGINIA	63	44	UNIVERSITY OF MINNESOTA	45
20	YALE UNIVERSITY	62	45	UNIVERSITY OF VIENNA	45
21	HARVARD MEDICAL SCHOOL	61	46	LEIDEN UNIVERSITY	43
22	NORTHWESTERN UNIVERSITY	61	47	UNIVERSITY OF CHICAGO	43
23	AARHUS UNIVERSITY	59	48	UNIVERSITY OF FLORIDA	43
24	OTTAWA HOSPITAL RESEARCH INSTITUTE	59	49	UNIVERSITY OF MANCHESTER	42
25	UTRECHT UNIVERSITY	59	50	UNIVERSITY OF WISCONSIN-MADISON	41

Table 6 provides a detailed overview of the most relevant affiliations in the field of open science research, showcasing the top 50 institutions by the number of articles published. Leading the list is the University of California with 255 articles, followed closely by the University of Toronto with 207 articles, and the University of Oxford with 197 articles. The table represents a diverse array of institutions, including renowned universities like Harvard Medical School, Yale University, and Stanford University, as well as specialised research centres like the Ottawa Hospital Research Institute. The presence of institutions from various countries, such as the University of Amsterdam, Aarhus University, and the University of Vienna, reflects the global reach and collaboration in open science research. The data underscores the significant contributions of these affiliations in advancing the field, highlighting the central role of academic and research institutions in fostering innovation and collaboration in open science.

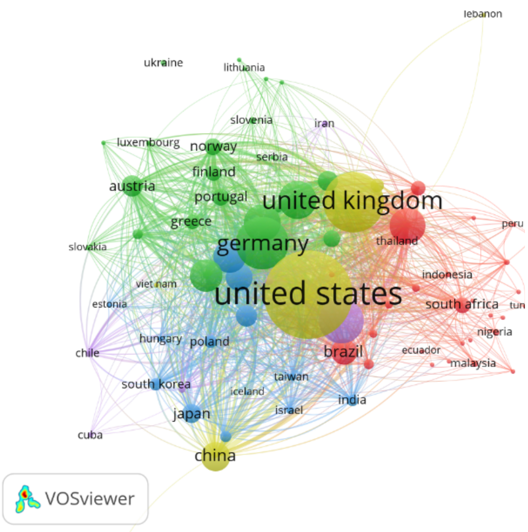


Figure 1. Collaboration Network by Countries

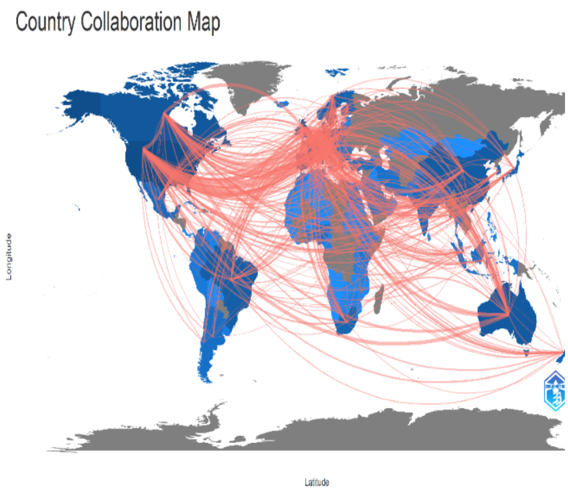


Figure 2. Countries' Collaboration World Map

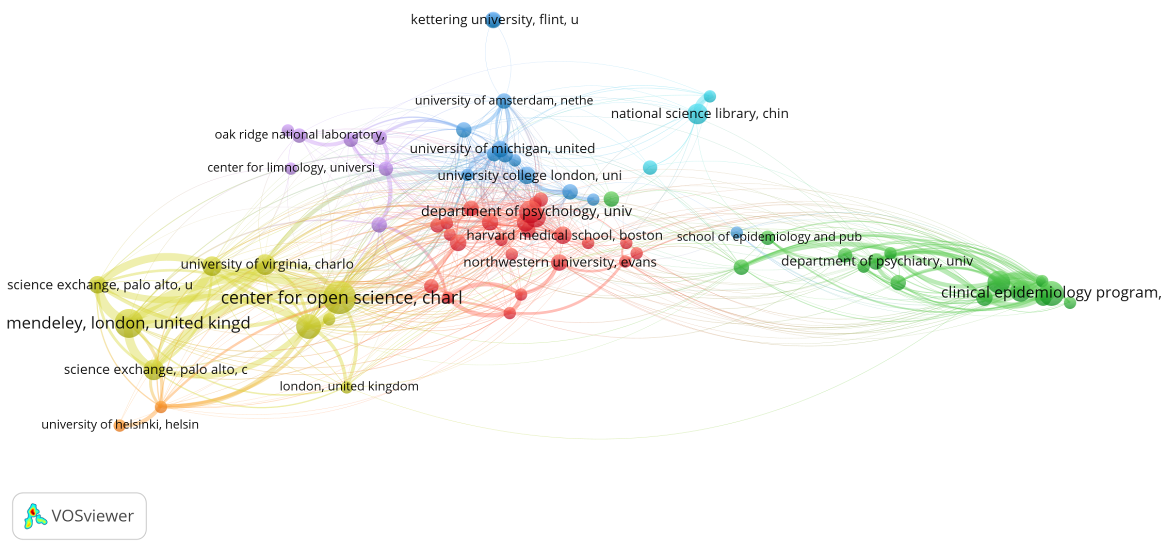


Fig. 11. Collaboration Network by Institutions.

Since the core focus of the open science movement was to encourage more transparent, collaborative, and inclusive scientific practices while also making scientific knowledge more accessible and verifiable (UNESCO, 2021). The collaboration network by countries in the context of bibliometric analysis of open science was also presented to illustrate a global synergy in scientific research. Key countries like the United States, the United Kingdom, and Germany emerge as central hubs, reflecting their significant investments and commitment to open science. The network also includes emerging economies like Brazil and India, showcasing the universal appeal of open science. The intricate connections between countries foster a collaborative environment that transcends geographical boundaries, promoting a cohesive approach to scientific discovery.

Similarly, the collaboration network by institutions reveals leading universities and research centers, such as the University of California and the University of Oxford, at the forefront of collaborative efforts. The network's diversity, including prestigious universities, research centers, and hospitals, enriches the research ecosystem. Strategic alliances between institutions lead to innovative outcomes and play a vital role in advancing the principles of open science. Together, these collaboration networks by countries and institutions underscore the shared commitment to knowledge, discovery, and the betterment of society, reflecting the boundless potential of collaborative research).

4.DISCUSSION

The main objective of this bibliometric analysis was to delve into the evolving landscape of Open Science through a comprehensive bibliometric study. Utilising a dataset from the Scopus online database, which is a robust information source encompassing scholarly literature across various disciplines, the study seeks to answer five fundamental questions related to the prevalence of topics, citations, leading journals, productive countries, and active institutions in open science research, and the analysis reveals the multifaceted nature of open science and its growing influence across various fields of research.

The topic analysis highlights the key focus and trending areas in open science, such as open access, reproducibility, data sharing, and collaboration. These themes reflect the core principles of open science, emphasising transparency, integrity, and accessibility in scientific research. The citation analysis offers insights into the influential works and authors that have shaped the field of open science over time, while journal analysis provides a view of the academic publishing landscape, including both traditional high-impact journals and innovative open-access platforms. Lastly, the country and institution analysis illustrate the global reach of open science, with contributions from a diverse range of countries and affiliations. The presence of both major research hubs and emerging contributors emphasise the universal appeal and applicability of open science principles.

Overall, the analysis underscores the significance of open science as a transformative approach to scientific research. By promoting openness, collaboration, and transparency, open science fosters innovation, accelerates knowledge dissemination, and enhances the credibility and reproducibility of research findings (Mahfooz & Roslina, 2021). The growing adoption of open science across various disciplines and regions reflects a positive shift towards a more inclusive and accountable scientific community.

As open science continues to evolve, ongoing research and analysis will be essential to track emerging trends, assess the impact of open practices, and guide future developments. The insights gained from this analysis contribute to a deeper understanding of the open science ecosystem and provide valuable guidance for researchers, policymakers, publishers, and other stakeholders committed to advancing the principles of open, collaborative, and transparent research.

5.CONCLUSIONS AND DIRECTION FOR FUTURE RESEARCH

The comprehensive analysis of open science research presented in this study offers valuable insights into the current landscape of the field. We have uncovered key trends, influential authors, prominent journals, and leading countries and institutions in open science research through topic analysis, citation analysis, journal analysis, and country and institution analysis. For the topic analysis, it was noticed that the most frequent occurrence words such as “open access,” “reproducibility,” “open data,” and “data sharing” highlight the core areas of focus in open science. These topics reflect the growing emphasis on transparency, collaboration, and accessibility in scientific research. While citation analysis, on the other hand, identifies the top-cited authors and documents and reveals the influential works that have shaped the field, providing a roadmap for future research directions and journal analysis evaluation of journals’ impact and relevance will offer a guide for researchers in selecting appropriate publication venues and assists librarians and publishers in enhancing their journals’ profiles, and the country and institution analysis present the geographical distribution of research output emphasises the importance of both national and international collaboration in shaping the open science landscape.

5.1. Direction for Future Research

Future research could delve deeper into emerging topics such as artificial intelligence, metadata, and ethics in open science, exploring their implications and applications across various disciplines. Investigating the collaboration dynamics between countries and institutions may uncover opportunities to foster greater international cooperation and interdisciplinary research. A critical examination of different open-access models and their impact on knowledge dissemination and accessibility could inform policy and practice in scientific publishing. Future studies could also assess how open science practices influence societal outcomes, such as public engagement with science, innovation, and education. Exploring the role of technology, including data management tools, repositories, and platforms, in facilitating open science would provide insights into the infrastructure needed to support this growing movement. Research into the barriers and enablers of open science in underrepresented regions could guide efforts to make open science truly global and inclusive. An in-depth analysis of how open science practices contributed to the rapid response to the COVID-19 pandemic could similarly offer lessons for future public health emergencies.

ACKNOWLEDGEMENT

The authors would like to thank the management and all staff of the Semantic Body of Knowledge and Technology Research Lab, IIUM for their contributions to the success of the research.

REFERENCES

Agarwal, A., Durairajanayagam, D., Tatagari, S., Esteves, S. C., Harlev, A., Henkel, R., Roychoudhury, S., Homa, S., Puchalt, N. G., & Ramasamy, R. (2016). Bibliometrics: tracking research impact by selecting the appropriate metrics. Asian Journal of Andrology, 18(2), 296.

Ahmed, M., & Othman, R. (2021). Readiness towards the implementation of open science initiatives in the Malaysian Comprehensive Public Universities. The Journal of Academic Librarianship, 47(5), 102368. https://doi.org/10.1016/j.acalib.2021.102368

Albalawi, R., Yeap, T. H., & Benyoucef, M. (2020). Using topic modeling methods for short-text data: A comparative analysis. Frontiers in Artificial Intelligence, 3, 42.

Aria, M., & Cuccurullo, C. (2017). bibliometrix: An R-tool for comprehensive science mapping analysis. Journal of Informetrics, 11(4), 959–975.

Banks, G. C., Field, J. G., Oswald, F. L., O'Boyle, E. H., Landis, R. S., Rupp, D. E., & Rogelberg, S. G. (2019). Answers to 18 questions about open science practices. Journal of Business and Psychology, 34(3), 257–270.

Burnham, J. F. (2006). Scopus database: a review. Biomedical Digital Libraries, 3(1), 1–8.

Cardenas-Gonzalez, M., & Alvarez-Buylla, E. R. (2020). The COVID-19 pandemic and paradigm change in global scientific research. MEDICC Review, 22(2), 14–18. https://doi.org/10.37757/MR2020.V22.N2.4

Castro, H., Pinto, N., Pereira, F., Ferreira, L., Ávila, P., Bastos, J., Putnik, G. D., & Cruz-Cunha, M. (2021). Cyber-Physical Systems using Open Design: An approach towards an Open Science Lab for Manufacturing. Procedia Computer Science, 196(2021), 381–388. https://doi.org/10.1016/j.procs.2021.12.027

Cathy, F. (2021, December 6). How the United Nations' new "open science framework" could speed up the pace of discovery. Science X Network. https://phys.org/news/2021-12-nations-science-framework-pace-discovery.html

Donthu, N., Kumar, S., Mukherjee, D., Pandey, N., & Lim, W. M. (2021). How to conduct a bibliometric analysis: An overview and guidelines. Journal of Business Research, 133, 285–296.

Enger, K. B. (2009). Using citation analysis to develop core book collections in academic libraries. Library & Information Science Research, 31(2), 107–112.

Essawy, B. T., Goodall, J. L., Voce, D., Morsy, M. M., Sadler, J. M., Choi, Y. D., Tarboton, D. G., & Malik, T. (2020). A taxonomy for reproducible and replicable research in environmental modelling. Environmental Modelling & Software, 134, 104753. https://doi.org/https://doi.org/10.1016/j.envsoft.2020.104753

Fischer, C., Pardos, Z. A., Baker, R. S., Williams, J. J., Smyth, P., Yu, R., Slater, S., Baker, R., & Warschauer, M. (2020). Mining big data in education: Affordances and challenges. Review of Research in Education, 44(1), 130–160.

FOSTER. (2020). What is Open Science? Introduction | Facilitate Open Science Training for European Research. FOSTER. https://www.fosteropenscience.eu/node/1420

Gallagher, R. V., Falster, D. S., Maitner, B. S., Salguero-Gómez, R., Vandvik, V., Pearse, W. D., Schneider, F. D., Kattge, J., Poelen, J. H., Madin, J. S., Ankenbrand, M. J., Penone, C., Feng, X., Adams, V. M., Alroy, J., Andrew, S. C., Balk, M. A., Bland, L. M., Boyle, B. L., … Enquist, B. J. (2020). Open Science principles for accelerating trait-based science across the Tree of Life. Nature Ecology and Evolution, 4(3), 294–303. https://doi.org/10.1038/s41559-020-1109-6

Garfield, E. (1994). The impact factor. Current Contents, 25(20), 3–7.

Go Fair. (2016). FAIR Principles - GO FAIR. Go-Fair. https://www.go-fair.org/fair-principles/

Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. Proceedings of the National Academy of Sciences, 102(46), 16569–16572.

Ibrahim, N. F., & Wang, X. (2019). A text analytics approach for online retailing service improvement: Evidence from Twitter. Decision Support Systems, 121, 37–50.

Kodonas, K., Fardi, A., Gogos, C., & Economides, N. (2021). Scientometric analysis of vital pulp therapy studies. International Endodontic Journal, 54(2), 220–230.

Kunze, K. N., Polce, E. M., Vadhera, A., Williams, B. T., Nwachukwu, B. U., Nho, S. J., & Chahla, J. (2020). What is the predictive ability and academic impact of the altmetrics score and social media attention? The American Journal of Sports Medicine, 48(5), 1056–1062.

Mahfooz, A., & Roslina, O. (2021). Promoting Open Science with Institutional Repositories in the Malaysian Comprehensive Public. Journal of Information Systems and Digital Technologies, 3(2), 11–28.

McAbee, S. T., Grubbs, J. B., & Zickar, M. J. (2018). Open Science Is Robust Science. Industrial and Organizational Psychology, 11(1), 54–61. https://doi.org/10.1017/iop.2017.85

Owan, V. J., & Owan, M. V. (2021). Complications connected to using the impact factor of journals for the assessment of researchers in higher education. Owan, VJ, & Owan, MV (2021). Complications Connected to Using the Impact Factor of Journals for the Assessment of Researchers in Higher Education. Mediterranean Journal of Social & Behavioral Research, 5(1), 13–21.

Perianes-Rodriguez, A., Waltman, L., & van Eck, N. J. (2016). Constructing bibliometric networks: A comparison between full and fractional counting. Journal of Informetrics, 10(4), 1178–1195. https://doi.org/https://doi.org/10.1016/j.joi.2016.10.006

Roldan-Valadez, E., Salazar-Ruiz, S. Y., Ibarra-Contreras, R., & Rios, C. (2019). Current concepts on bibliometrics: a brief review about impact factor, Eigenfactor score, CiteScore, SCImago Journal Rank, Source-Normalised Impact per Paper, H-index, and alternative metrics. Irish Journal of Medical Science (1971-), 188, 939–951.

Sivertsen, G., Rousseau, R., & Zhang, L. (2019). Measuring scientific contributions with modified fractional counting. Journal of Informetrics, 13(2), 679–694.

Spitschan, M., Schmidt, M. H., & Blume, C. (2021). Principles of open, transparent and reproducible science in author guidelines of sleep research and chronobiology journals. Wellcome Open Research, 5. https://doi.org/10.12688/wellcomeopenres.16111.2

Suber, P. (2010). A very brief introduction to open access.

Tennant, J. P., Agrawal, R., Baždarić, K., Brassard, D., Crick, T., Dunleavy, D. J., Rhys Evans, T., Gardner, N., Gonzalez-Marquez, M., & Graziotin, D. (2020). A tale of two'opens': intersections between Free and Open-Source Software and Open Scholarship.

Tenopir, C., Rice, N. M., Allard, S., Baird, L., Borycz, J., Christian, L., Grant, B., Olendorf, R., & Sandusky, R. J. (2020). Data sharing, management, use, and reuse: Practices and perceptions of scientists worldwide. PLoS ONE, 15(3), 1–26. https://doi.org/10.1371/journal.pone.0229003

Tomaszewski, R. (2023). Visibility, impact, and applications of bibliometric software tools through citation analysis. Scientometrics, 1–22.

UNESCO. (2021). UNESCO Recommendation on Open Science.