

Global Research Trends in Research Data Management (Rdm) – A Scientometric view

Naseema S

Senior Research Fellow, Department of Library and
Information Science, Pondicherry University,
Puducherry, Kalapet, India.

naseemaphd3@gmail.com

ORCID iD: <https://orcid.org/0000-0003-1952-5934>

R Sevukan

Professor, Department of Library and Information
Science, Pondicherry University,
Puducherry, Kalapet, India.

Corresponding Author: rsevukan.lis@pondiuni.edu.in

ORCID iD: <https://orcid.org/0000-0001-8648-3952>

Received: 21 August 2021

Accepted: 14 November 2021

Abstract

The study focuses on Research Data Management (RDM), aiming to demonstrate how RDM is evolving globally. It provides a systematic mapping of the current literature to aid in identifying core coverage and reflecting on potential trends using the scientometric technique. For detailed thematic analysis, performance analysis, network representation, science mapping, and scientific collaboration, this study used bibliometric tools such as the R package Biblioshiny, SciNetPy, and VOSviewer. Two premier indexing databases, SCOPUS and Web of Science, which extensively cover RDM literature, are appropriately considered. Furthermore, all retrieved documents are refined by language, document type, and irrelevant keywords for practical analysis. This study covers the Research Data Management (RDM) literature from 1926 to 2020, with 6263 documents published from 1666 sources and 15,545 authors. The most common domains discovered in RDM research are Computer Science, Library and Information Science. RDM is primarily familiar and cultivated in nations such as the United States, China, the United Kingdom, Canada, Germany, and Australia are pioneers in RDM research. The findings are vital for researchers working on RDM projects and policy formulation. This research helps to identify the literature's strengths and potential gaps. This secondary data contributes significantly to the scientific landscape in scientific production, network architecture, source clustering, and international collaboration, the evolution of dominant subjects and countries, and science mapping of productive word frequency.

Keywords: Data Management, Research Data, Scientific Collaboration, RStudio, Science Mapping, Research Life Cycle, Data Centre.

Introduction

Research data is a byproduct of the research lifecycle; it is used and produced in the form of publications. However, the scientific community considers it like a skeleton in the cupboard. Many researchers defined Research Data Management (RDM) as "data are all complex data entities from observations, experiments, simulations, models, and higher order assemblies, along with the associated documentation needed to describe and interpret the data" (Cyberinfrastructure Council, 2007). The process of controlling and organizing the generated

data during a research project covers the entire data lifecycle, which includes the planning of the investigation, conducting the investigation, storing and backing up of the data as it is created, and preserving the data long-term after the research investigation has concluded (Penn State University Libraries, 2014). The collection and processing of data in the research process are usually referred to as research data, and it varies from discipline to discipline. Statistics, experimental results, observational data, and interview recordings are all examples of research data (Anilkumar, 2018). Research data is the "black box" of any research work or project (Cox & Tam, 2018). It is essential to safeguard research data, and the reuse of research data by research organizations is in its early stages. Many institutions are unsure how well the research projects back up their data after completion.

A few studies from India have been reported in this domain. Bunkar and Bhatt (2020) revealed that data sharing, accessibility, and the library's involvement in organizing and preserving research data are essential. The rapid growth of scientific publications fails to manage the research data as "it requires advanced data management skills, techniques, and tools that researchers often require assistance with" (Joo & Peters, 2020). Gunjal and Gaitanou (2017) proposed an RDM framework using research products implemented in Central Library at NIT Rourkela in Odisha, India. Tripathi, Shukla, and Sonker (2017) surveyed the central university libraries in India and found that it is at the initial stage of implementing RDM services. Tripathi, Chand and Sonkar (2017) demonstrated uncertainty about data sharing, problems with data storage, and significant interest in library support with these aspects of data management. Due to the aforementioned, this study seeks to assess, using scientometric methodologies, the global scenario of Research Data Management (RDM) in terms of author networks, significant sources for RDM publications, core areas focused on in RDM research, main nations actively interested in RDM research, and scientific network of various countries and institutions involved in RDM research.

Review of Literature

It has been observed in the global scenario of RDM research in the recent past. Chiware (2020) explored the African academic and research library's data management service using qualitative content analysis and reported that few countries are in the formative stages. Research Data Management (RDM) policies have emerged, cyber and data infrastructure are being developed, and limited data librarianship courses are being taught. Mushi, Pienaar and Deventer (2020) studied RDM service at the University of Dodoma library, Tanzania. Chawinga and Zinn (2020b) investigated the status of RDM at two public university libraries in Malawi and identified that RDM is a new and emerging concept. Choi and Lee (2020) surveyed the RDM status of science and technology research institutes in Korea. They suggested that creating an open science ecosystem by creating, managing, and utilizing research data in a FAIR (Findable, Accessible, Interoperable, Reuse) will be the foundation for the next generation of scientific research. Chawinga and Zinn (2020a) explored the status of RDM at a medical school in Africa, resulting in the health research community's better understanding of RDM perspectives in developing countries. Roche et al. (2020) described Canada's federal initiatives promoting open government data and environmental science datasets, summarised the RDM challenges, and planned to modernize the open data approach and best practices for data discoverability, access, and reuse.

Redkina (2019) explored the current trends in RDM and concluded that it is necessary to

develop service management and create institutional and national research data management services. Ashiq, Usmani and Naeem (2020) reviewed the RDM practices and services using a systematic literature review, and a meta-analysis found that RDM is gradually gaining importance among researchers and academic libraries. Still, it's poorly practiced in developing countries. Bangani and Moyo (2019) investigated the public universities in South Africa and found that researchers preferred to use data produced by others but were less keen on sharing their data; A similar study was conducted in other countries such as the United Kingdom, France, and Turkey but the Researchers believe that circumstances in the developed world may differ with the South African research environment. Kun, Tong, and Xiaodan (2019) explored the application methods of big data technology in the military and scientific research data management and laid a foundation for the technical plan and construction of a scientific research big data platform for military enterprises. Marlina and Purwandari (2019) synthesized the RDM Strengths, Weaknesses, Opportunities, and Threats (SWOT), which were mapped into the SWOT matrix using Systematic Literature Review (SLR) to develop strategies to implement RDM services in Indonesia. Cho (2019) extracted metadata from re3data and analyzed the current status of research data repositories in Asian countries, found that 70 percent of the Asian-country repositories are those where licenses or policies are declared but not granted as permanent identifiers and international level certification, Life sciences and natural sciences are center clustering domain of re3data repositories.

Koltay (2017) examined the role of Research 2.0 and Research Data Services in academic and research libraries. Liu and Ding (2016) investigated the feasibility of university libraries in providing RDM services without any supporting policies from governments or funding agencies. Onyancha (2016) explored the status of research data sharing among researchers in Sub-Saharan Africa (SSA) and internationally using Data Citation Index (DCI) and found that research institutions and universities are the major contributors to research data, and there is a significant correlation between the publication of data and research articles. Surkis and Read (2015) highlighted the importance of RDM service to be rendered by librarians with little or no experience in the RDM domain. They suggested that librarians acquire new skills for providing services on RDM. Corral, Kennan, and Afzal (2013) investigated the library support for Research Data Management (RDM) in Australia, New Zealand, Ireland, and the United Kingdom and found a significant difference between countries and librarians' needs for a multilayered understanding of the research environment.

Research Questions

- Q1: What is the worldwide trend in scientific publishing in Research Data Management (RDM)?
- Q2: How is RDM literature growing, and in which direction is it trailing or remaining unexplored?
- Q3: Who are the major contributors (country) globally in the field of RDM?

Objectives of the Study

- To analyze the evolution of scientific production and the author network in RDM research.
- To highlight the most important relevant sources for RDM publications.
- To find out the core areas focussed in RDM research

- To identify the significant countries actively involved in RDM research.
- To investigate the scientific network of various countries and institutions involved in RDM research.

Materials and Methods

The data were retrieved from the Web of Science (WoS) Core Collection and SCOPUS. The search query was formulated with appropriate keywords after careful examination to obtain precise results. The query statements explored in these two databases are as follows:

SCOPUS - TITLE-ABS-KEY (ALL ("Research Data Management*" OR "RDM" OR "Research Data Set*" OR "Research Data Preservation*" OR "Research Data Archiv*" OR "Institutional Repositor*" OR "Research Information Polic*" OR "Intellectual Cooperation*" OR "Research Moral and ethical aspect*" OR "Research Information Service*" OR "Data Center"))

WEB OF SCIENCE – ALL FIELD ("Research Data Management*" OR "RDM" OR "Research Data Set*" OR "Research Data Preservation*" OR "Research Data Archiv*" OR "Institutional Repositor*" OR "Research Information Polic*" OR "Intellectual Cooperation*" OR "Research Moral and ethical aspect*" OR "Research Information Service*" OR "Data Center"))

Thus, the data were retrieved from WoS in Win, UTF-8, and SCOPUS in .csv excel formats. Simple Keywords and Boolean operators were used to retrieve the results effectively (Janik, Ryszko, and Szafraniec, 2020). For 94 years between 1926 and 2020, SCOPUS returned 44,218 documents, and WoS returned 21,191 documents. It was noticed that a few records were overlapping, which had to be filtered to make the data complete without overlapping. To perform this function, ScientoPy software was used to normalize the unique records to reduce data redundancy is one of the study's distinct features. The data were collected during April 2021. As shown in Figure 1, the total number of documents collected from WoS and SCOPUS comes to 6853 records with a break up of 2539 and 4314, respectively. Of these, 590 (WoS=4, SCOPUS=586) records were overlapped, and the final number of records after filtration came down to 6263 (WoS=2535, SCOPUS=3728) records for further analysis.

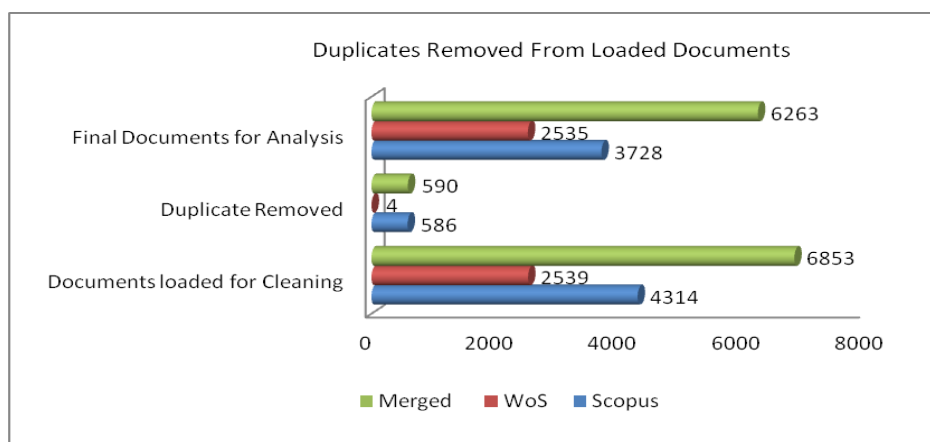


Figure 1: Process of removing records overlapping (Source: *Figure extracted from ScientoPy)

In addition, some software tools such as R version 4.0.2 (Bibliometrix package) to perform comprehensive science mapping discovery (R Core Team, 2020); ScientoPy (v 2.0.3) for pre-

processing to improve dataset precision and readability (Ruiz-Rosero Ramirez-Gonzalez, Williams, Liu, Khanna & Pisharody, 2017); Biblioshiny (v 2.0) for quantitative research in scientometric and bibliometric methods of analysis (Bibliometrix R Package, n.d.); and VOS-viewer (v 1.6.16) for constructing and visualizing bibliometric networks (Alonso, Castiello & Mencar, 2018) were employed in the study.

Results

RDM Research Profile Summary

As extracted from Biblioshiny, Table 1 presents the complete research profile of research data management (RDM) for the period spanning between 1926 and 2020. It is clear from the result that a total of 6263 publications scattered over 1666 sources were found in SCOPUS and WoS, with an average of 6.23 papers per year. Concerning the author profile, there is a record of contribution from 15545 authors with 34019 appearances. Of which, 1162 authors contributed as a single author while 14533 contributed as collaborative authors. Thus, the average number of documents per author comes to 0.403; the average number of authors per document is 2.48; the average number of co-authors per document is 5.43, and the collaborative index is 2.85. The research in the field of RDM is found to be reasonably collaborative. As stated earlier, eleven (11) query phrases were formulated to retrieve the most relevant documents. On the other hand, the impact of RDM witnessed 13.17 as an average number of citations per document and 1.935 as an average number per year for a record.

Table 1

Scientometric Profile of Research Data Management (RDM)

| Parameters | Description | Values* |
|-------------------------------------|--|-------------|
| Time span | Period of Coverage | 1926 – 2020 |
| Documents | Total number of documents | 6263 |
| Year wise Average | The average number of publications per year | 6.23 |
| Sources | Distribution of publication sources | 1666 |
| Authors | Total number of authors | 15545 |
| Author Appearances | Frequency distribution of authors | 34019 |
| Single-authored documents | No. of single authors | 1162 |
| Authors of multi-authored documents | No of multi-authors | 14533 |
| Documents per author | Average no. of documents per author | 0.403 |
| Authors per Document | Average no. of authors per document | 2.48 |
| Co-Authors per Document | Average no. of co-authors per document | 5.43 |
| Collaboration Index | An index of single Vs. multi-authored publications | 2.85 |
| Keywords Plus (I.D.) | No. of terms frequently appeared in the title of an article's references | 9301 |
| Author's Keywords (D.E.) | Total number of keywords | 12313 |
| Average citations per document | No. of citations received per document | 13.17 |
| Average citations per year | An average of citations per year of a document | 1.935 |

*Values extracted from Biblioshiny

Research Trends in RDM

Figurer 2 indicates the annual growth of research publications in Research Data

Management from 1926 to 2020 based on WoS and Scopus. The results show that the number of publications has gradually increased from beginning to end, except for 2020. In comparison, publication growth from 1926 to 2006 was comparatively low. From 2008 onwards, the number of publications rose to over a hundred.

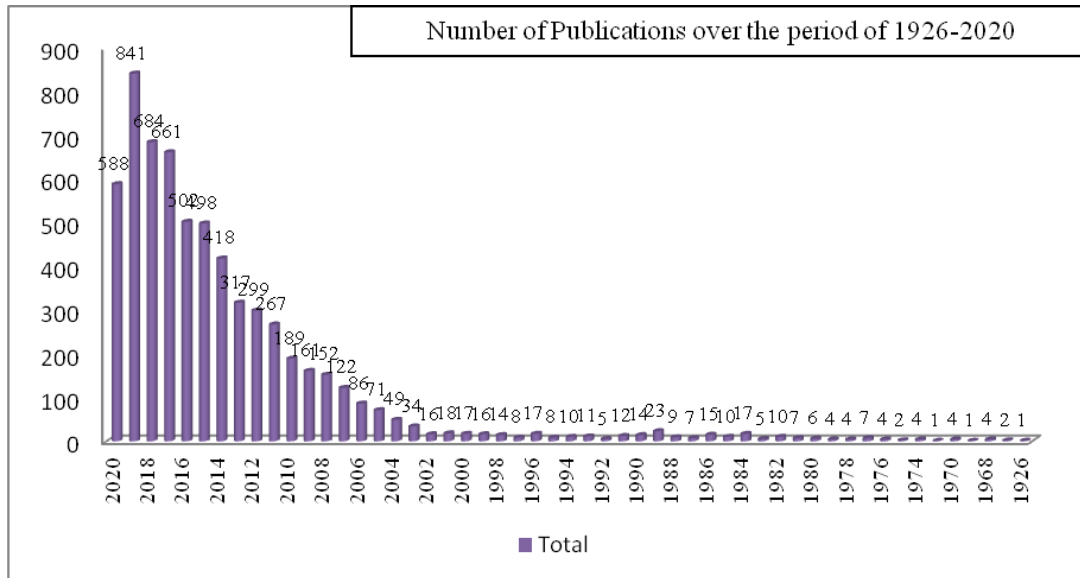


Figure 2: Annual Scientific Production (Source: *Figure extracted from Biblioshiny)

Mapping of Top Countries, Subject Areas, and Source

Figure 3 depicts the scientific mapping of RDM literature to visualize the interrelationships between countries that contributed the most in subject domains and publication sources. The major contributing countries appear on the left side of Figure 3, while the leading research areas and predominant sources of publications appear on the middle and right sides, respectively. The result indicates that the USA ranks first in order, followed by India, China, the United Kingdom, Germany, etc.

Interestingly, most of the contributions made by Indian researchers in the field of RDM were in the top-ranked vital areas. In contrast, a minor concentration was noticed in the areas such as virtualization, data centers, energy efficiency, and university libraries. Out of various sources of publications preferred by the RDM researchers worldwide, it was observed that the journals such as the International Journal of Science, Library Hi Tech, New Review of Academic Librarianship, Electronic Library, and Future Generation Computer Systems covered most of the publications in RDM.

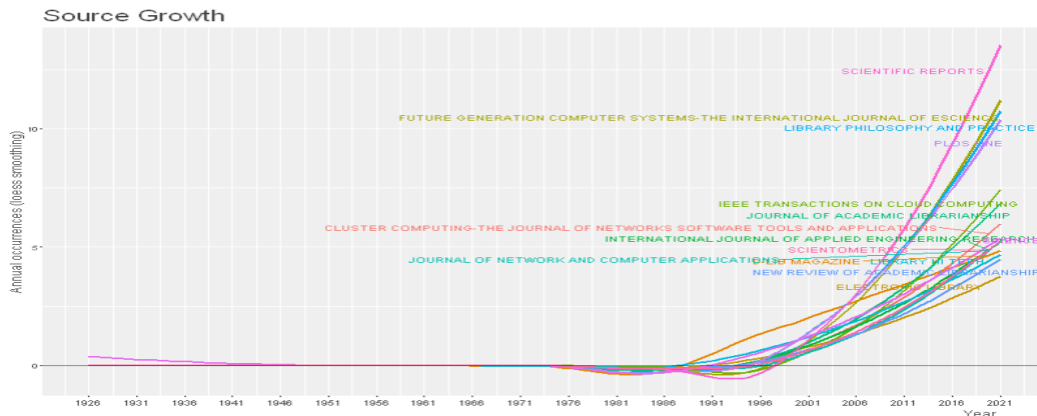


Figure 4B: Source Growth (Source:*Figure extracted from Biblioshiny)

Productive Subjects in RDM Research

Research Data management is widespread in all scientific domains of knowledge. As depicted in Figure 5, the analysis mainly identifies impacted areas and focuses on this research domain. The top and leading research domains were computer science, library and information science, science and technology, telecommunications, educational research, management and operational research, and management and operational research. The other topics listed in Figure 5 were less prevalent in RDM research.

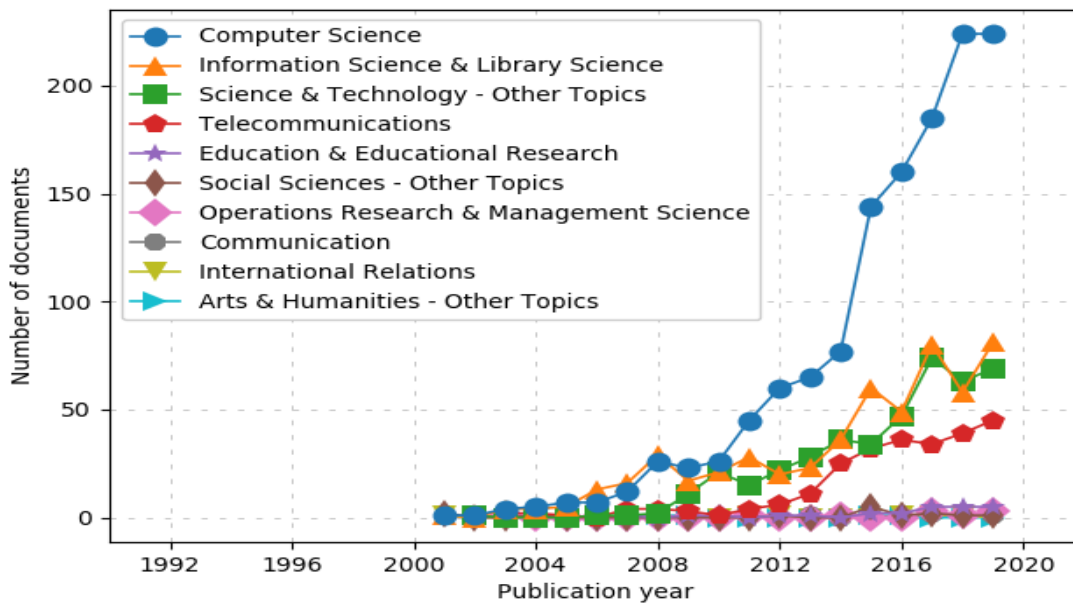


Figure 5: Productive Subjects in RDM Research (Source:*Figure extracted from ScientoPy)

Most Productive Authors

Based on the data of RDM retrieved from SCOPUS and Web of Science, it was observed that 15545 authors and 34019 collaborators appeared in 6263 papers. The author analysis is depicted in Figures 6A and 6B, which helps identify the most significant contributions in RDM publications. The top twenty authors are represented in Figure 6A. Buyya, R. is ranked top, reporting 23 publications with an h-index of 14, followed by Li, K.Q. with 15 publications (h-index=10). Cox, A.M. and Khan, S.U. have ten publications with an h-index of 8. Li, Y., has

published 19 papers. However, his h-index is 7 points lower than that of previous contributors. Calheiros, R.N. has eight publications with a seven h-index, while Abrizah, A. and Xia, J.F. have nine publications with a seven h-index. Other authors with lower h-index include Kim, J., Chen, Y., Zhang, Y., Jia, X.H., Schopfel J., Fan, J.X., Wang, X., and Wolschin G.

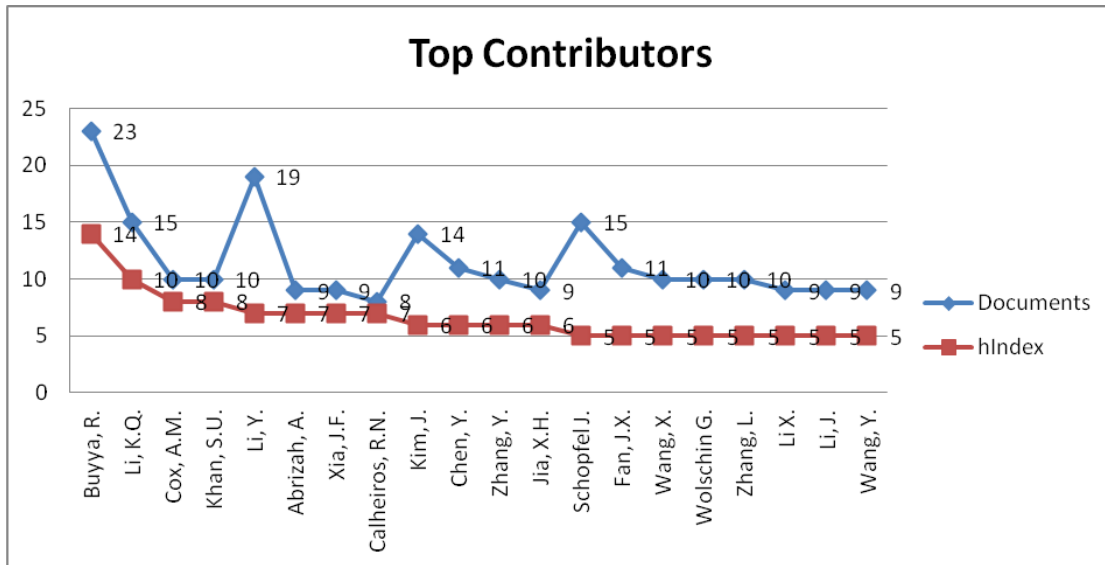


Figure 6A: Most Productive Authors (Source:*Figure extracted from ScientoPy)

Figure 6B compared the co-authorship network with top contributors, representing author collaboration by recognizing similar colors as clusters. Top contributors and co-authorship network's prominent scholars are both identical.

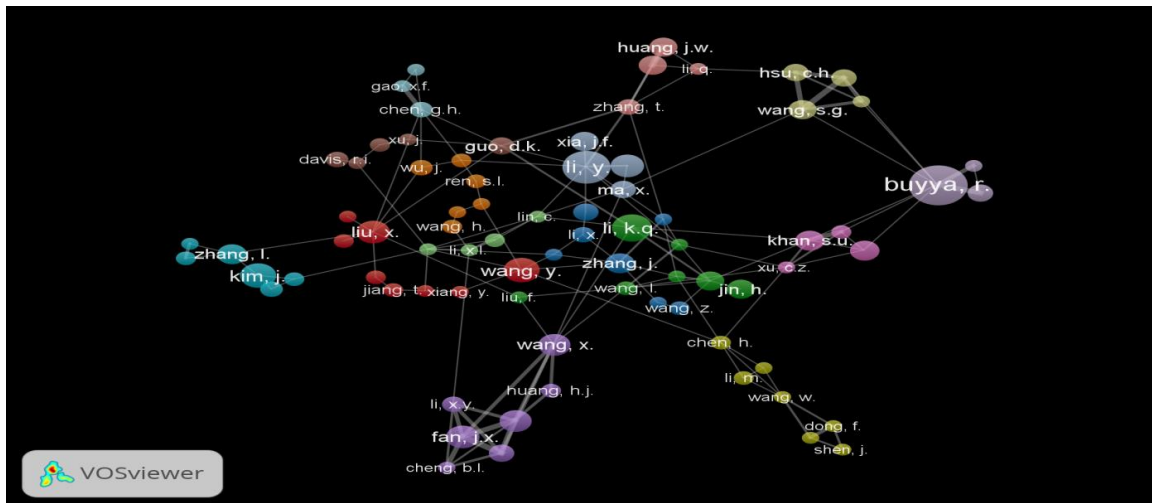


Figure 6B: Co-authorship Analysis (Source:*Figure extracted from VOSviewer)

Corresponding Author's Collaboration Pattern

According to the analysis, two categories were used for the analysis of the corresponding author's country: Single Country Publication (SCP) which refers to the authors collaborating within the country, and Multiple Country Publication (MCP), which relates to the authors collaborating within and outside the country. The ratio of SCP to MCP tends to vary substantially from country to country.

Figure 7A shows the corresponding authors' countries, which are the United States (SCP=543, MCP=61), United Kingdom (SCP=105, MCP=18), India (SCP=90, MCP=4), China (SCP=74, MCP=14), Germany (SCP=56, MCP=6), and both Korea and Australia (SCP=47,46 & MCP=10), followed by other corresponding contributions in this domain. Similarly, Figure 7B demonstrates the most often cited countries, including the United States, United Kingdom, Spain, Canada, Germany, and so on, with a significant rise compared to the corresponding author's country. The investigation identifies that more contributing countries significantly influence Single Country Publication (SCP) and Multiple Country Publication (MCP).

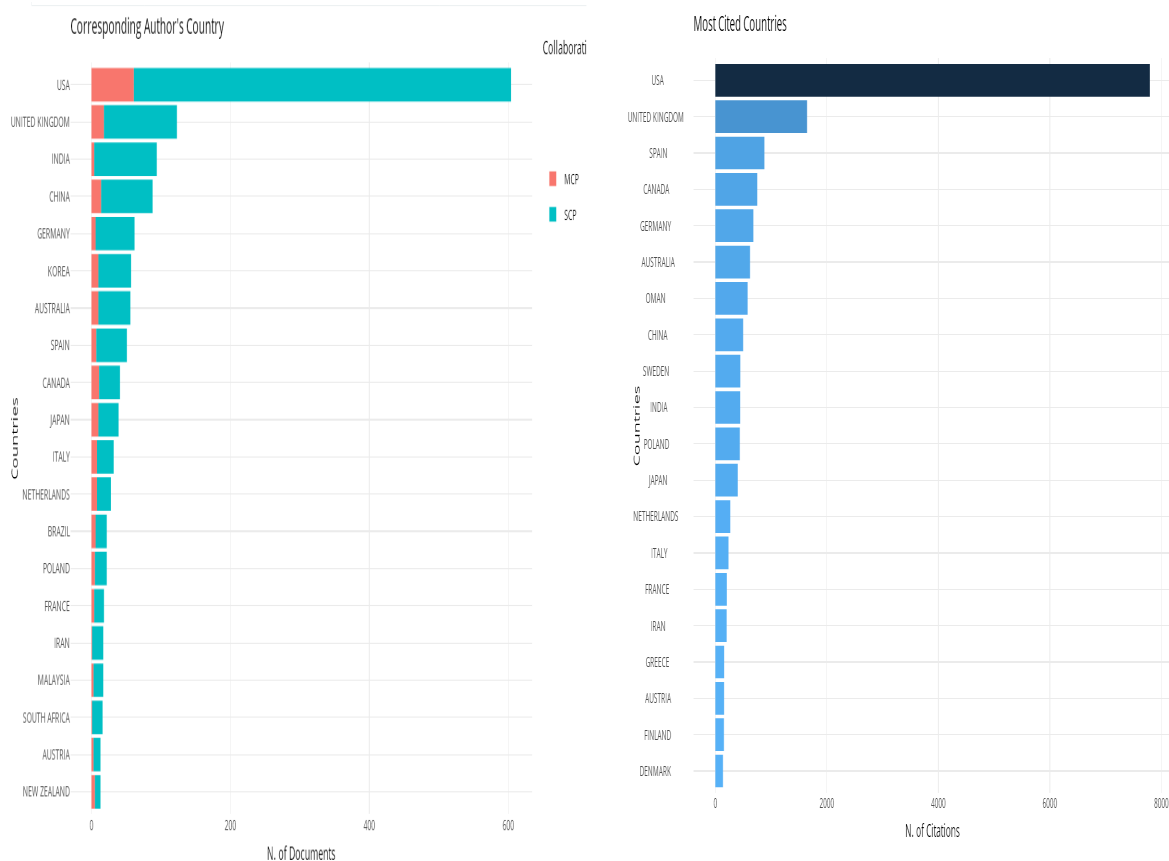


Figure 7A: Corresponding Author's Countries

Figure 7B: Most Cited Countries

(Source: *Figure extracted from Biblioshiny)

Mapping of Productive Terms and Word Cloud

Figures 8A and 8B compare the keyword patterns and highlight significant keywords and their network mapping, which helps identify related concepts in the study field. Figure 8A provides a word cloud of crucial topics that includes data centers, information management, digital storage, institutional repositories, and other essential things that are larger. Similarly, in Figure 8B, the network of frequent words is shown in many colors, each reflecting the same concepts as a related cluster connected to the core concept. This can assist in determining the primary area of network mapping and the significant clusters of words inside it. The core clusters include data centers (red), institutional repositories (purple), open access (orange),



Figure 9: Collaboration network of Countries (Source: *Figure extracted from Biblioshiny)

Table 2 determines the proportion of contributions between the United States and India, which is 50, followed by the United Kingdom (38), China (31), Georgia (22), Canada (19), Australia, Germany, Mexico (15), and Japan (11). Similarly, the combined contribution of the United Kingdom and Australia comes to 25, while the collaborative contribution of Germany and the United Kingdom is 16. China and the United Kingdom have collaborated with 11 contributions each.

Table 2

Mapping of Collaboration among Countries

| Rank | From | To | Frequency |
|------|----------------|----------------------------|-----------|
| 1 | USA | India | 50 |
| 2 | USA | United Kingdom | 38 |
| 3 | USA | China | 31 |
| 4 | United Kingdom | Australia | 25 |
| 5 | USA | Georgia | 22 |
| 6 | USA | Canada | 19 |
| 7 | Germany | United Kingdom | 16 |
| 8 | USA | Australia, Germany, Mexico | 15 |
| 9 | China | Hong Kong | 12 |
| 10 | Oman | Romania | 12 |
| 11 | China | United Kingdom | 11 |
| 12 | Serbia | France | 11 |
| 13 | United Kingdom | Canada | 11 |
| 14 | USA | Japan | 11 |

Growth Trend of Countries' Contribution to RDM

As shown in Figure10, the top ten trending countries' publications are analyzed based on the total number of papers published during the year. This evolution graph assists in identifying the leading countries that have published the most articles, especially compared to

the previous year's trends. On the left, the cumulative number of publications (on a logarithmic scale) is shown against the year of publication. The progression starts with the publication year on the X-axis, while the Y-axis indicates the number of documents published by each country. On the right side plot, the X-axis shows the proportion of documents published in the previous years (2018-2019), while the Y-axis represents the average number of documents published every year (2018-2019) (Ruiz-Rosero et al., 2017). Analysis indicates which countries have the highest average growth rate and the most significant proportion of recent articles published. The emerging countries in RDM research and development are the United States, China, and India, in third place in collaborative contribution. Germany, Australia, Canada, Spain, South Korea, and Japan were in the last place in joint contributions to RDM research.

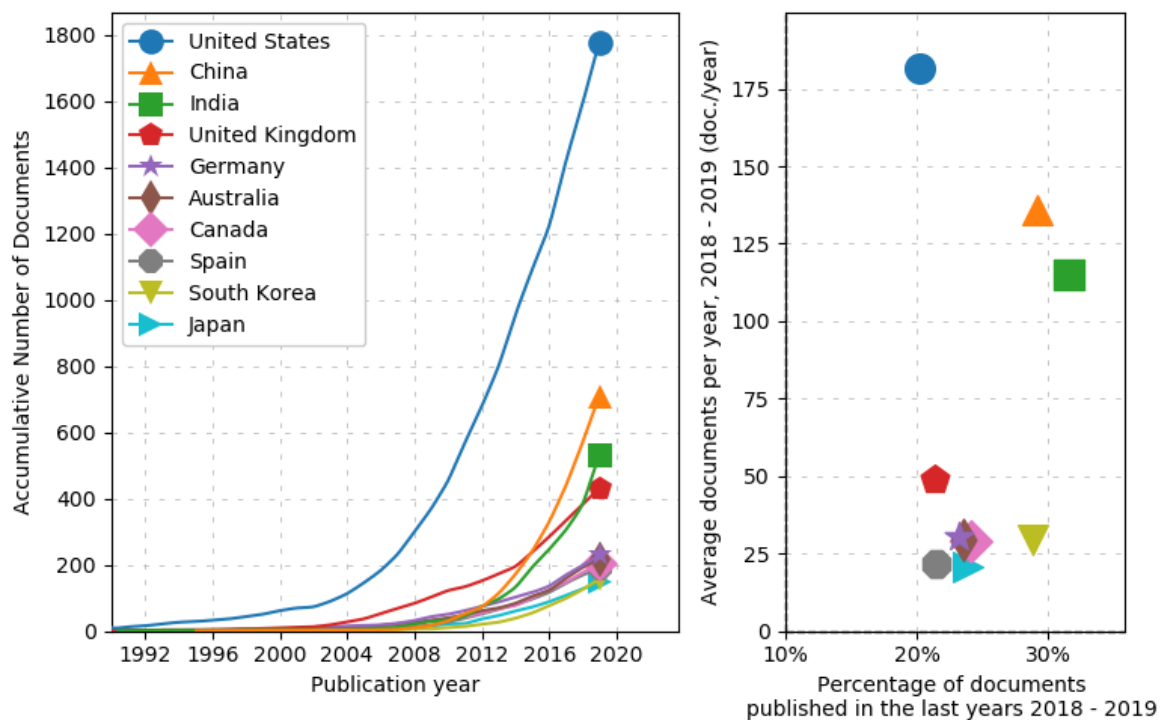


Figure 10: Emerging countries in collaborative research (Source: *Figure extracted from ScientoPy)

Global Institution's Contribution

This study aims to identify institutions participating in Research Data Management Practices and Services. Figure 11 shows the overall number of papers published and the proportion of documents published in the most recent years (2018 to 2019) to identify the top institutions globally. It was observed that institutions from China and United States have more influence than institutions from other countries. It is evident that the University of Chinese Academy of Sciences (China) has contributed 50 percent of publications during 2018-2019, followed by The National University of Defence Technology (China) (43 percent), the University of Melbourne-Australia (42 percent), the Duke University-United States and Beijing University of Posts and Telecommunications-China (41 percent), the University of Illinois-United States, and Shanghai Jiao Tong University-China (39 percent), the Technical University of Berlin-Germany (35 percent), and the Chinese Academy of Sciences-China (31 percent). All other universities have contributed a few publications related to RDM Practices and Services.

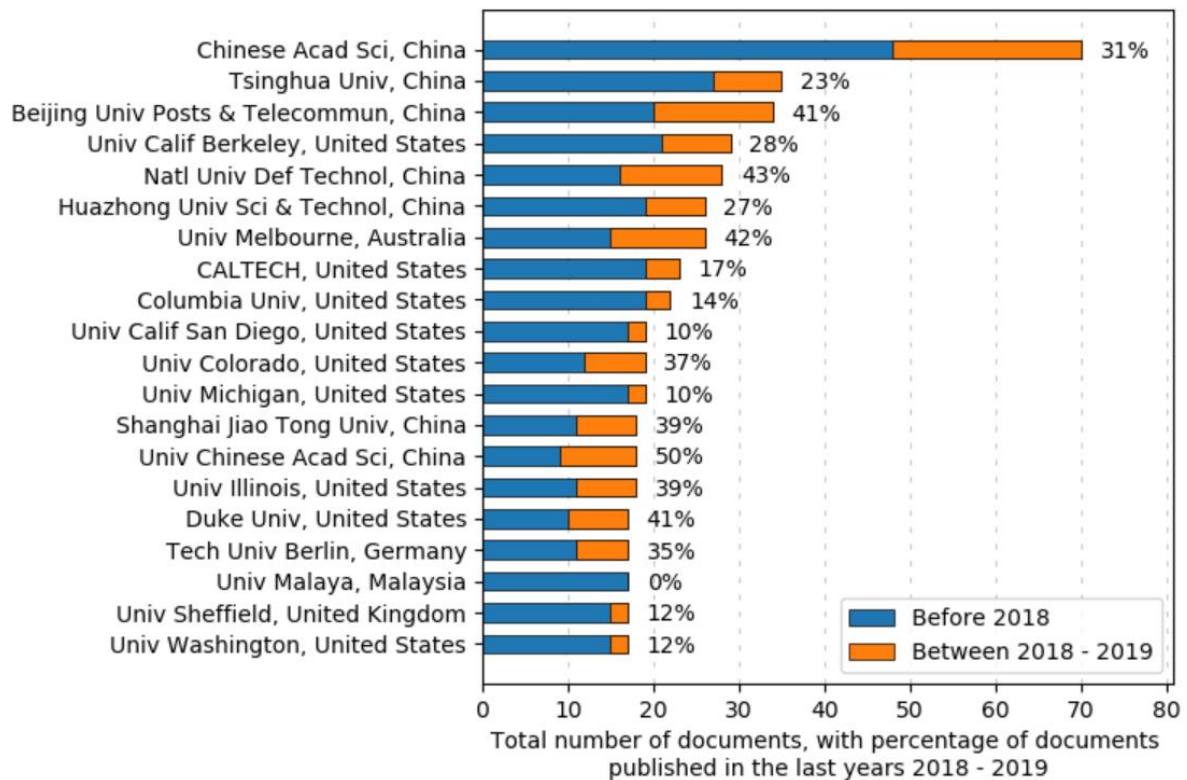


Figure 11: Global Institution's Collaboration (Source: *Figure extracted from ScientoPy)

Discussion

Using secondary data, this study critically reviewed the RDM literature. In recent years, the number of publications has skyrocketed, reaching an all-time high of 841 in 2018 received publications related to this study. In the last ten years, scientific production in RDM has quadrupled. It represents the advancements of RDM practices and services carried out by various subjects, with Computer Science and Library & Information Science leading the way globally, followed by Science & Technology, Educational Research, Social Science, Management & Operations Research, Communication, International Relations, and finally Arts & Humanities. Payal, Awasthi and Tripathi (2019) reviewed that library professionals must update and up-skill themselves with new trends, tools, and techniques to provide RDM services. Anilkumar (2018) observed that libraries' role in data management was still at a very early stage of development in India. Finding the exact competencies that library professionals will need to train, operate, manage, and deliver RDM services shortly is vital. Professionals in the library sector can collect data more effectively than other domain experts.

Surprisingly, India is in the second position in the research productivity of RDM. The primary research topics of RDM researchers were discovered to be cloud storage, open access, institutional repositories, academic libraries, and scholarly communication, which significantly impact India's top ranking. In a broad sense, the concept of RDM dates back to 1926, and fundamental research advancements are coming from various countries published in the current decade. Metadata, Institutional Repository, Institutional Policy, and Copyright were all needed to facilitate RDM development. Late and Kekalainen (2020) studied the purpose and user profile of dataset downloaders of the Finnish social science archive. They revealed that quantitative datasets are significantly utilized, and all stockholders reusing research data would

increase shortly.

The contribution of the corresponding author's country witnessed that single contributing publications (SCP) outnumbered multiple contributing countries (MCP). In the same way, the countries of the corresponding authors were identified and then analyzed to comprehend international collaboration on Research Data Management (Liu, Peng, Li & Cao, 2019). More importantly, cross-country network makes it easier to get a sense of scientific collaboration at various levels of aggregation (Fusco, Marsilio & Guglielmetti, 2020; Biancone, Saiti, Petricean & Chmet, 2020). By the way, it demonstrates that the development of RDM projects is more concentrated on inter-country advancements than intra-country contributions. Significantly most involved SCPs include the United States, the United Kingdom, India, China, Korea, Australia, and Spain. A few countries, such as the United States, the United Kingdom, China, and Canada, are also participating in MCP. Subsequently, the most often cited countries are the United States, the United Kingdom, Spain, Canada, Germany, Australia, Oman, China, Sweden, India, Poland, Japan, the Netherlands, and Italy, among others. A few studies explored RDM services. Liu, Zotoo, and Su (2020) identified the existing RDM policies and differentiated between the three developed nations (USA, U.K., and Australia) using a content analysis approach and ascertained the areas of access, retention, sharing, storage, and ownership as common underlying facts. Data Management Plan (DMP) is essential to enable the policymakers in developing countries to draw empirical evidence from the developed countries on RDM. This will form the basis for policy directions. Herbst, et al. (2015), complementary to this study, developed a network using low and middle-income countries' longitudinal data framed as The International Network for the Demographic Evaluation of Populations and Their Health (INDEPTH) to contribute to and share high-quality datasets with the scientific community and health policymakers.

Institutional Repository (I.R.), Open Access, Data Centre, Cloud Computing, Data Management, Metadata, Digital Libraries, Big Data, and Scholarly Communication were recognized as relevant areas that aid in the development of Research Data Management in this study. The keyword mapping and word cloud are used to determine these keywords. Interestingly, the number of authors' keywords (12,313) indexed in databases was almost double the total number of documents. Similarly, 9301 keywords occurred commonly in articles' titles of the cited references, which are smaller than the Authors' keywords (Secinaro, Brescia, Calandra & Biancone, 2020). This study aided in the discovery of fundamental ideas required for the development of RDM services.

Another essential point that Saeed and Ali (2019) found is that social science scholars are more willing to share than life science. The primary two domains involved in this research are computer science and library & information science, so the majority of the publications are from these two domains, such as predominant journals related to library science such as Library Hi-Tech, Serials Review, Electronic Library, Journals Of Academic Librarianship, Library Philosophy & Practices, Scientometric, New Review Of Academic Librarianship, D-Lib Magazine, Gray Journal, a data science journal. International Journal of Applied Engineering Research, Future Generation Computer System, IEEE Transactions on Network and Service Management, Cluster Computing, IEEE Transactions on Cloud Computing, ACM Sigcomm Computer Communication Review, and finally, Journal of Network and Computer Applications are prominent journals publications related to RDM by computer science and technology. In the same way, key publications have been reported in Computer Science and Library and

Information Science Journals (Liao & Ma, 2018).

The growth of publications and recent developments were determined with the help of this investigation. For each country, the analysis is based on comprehensive documents with a percentage of documents published in the previous year. According to the publications, the highest priority institutes are from China, the United States, Germany, Malaysia, the United Kingdom, Spain, Iran, and South Korea. It demonstrates that the countries mentioned earlier are primarily concerned with preserving and sharing research data management. According to the statistics, none of the Indian institutions are especially active in this study compared to worldwide universities. It demonstrates the critical need for India to work in this sector on a large scale and with other nations. As the global economy improves, developing and underdeveloped countries that spend much money on research and publish many papers must lead the way in storing, preserving, sharing, and reusing research data. Developed countries have an established practice of RDM usage and services, which can be investigated to learn about and improve future processes if any issues are discovered and to design a common framework for all research data management practitioners. To that end, it is necessary to examine user profiles and find statistics about the users who use the deposited data, such as what purpose it is being used for, how efficiently it is being used, who all the users are, and what credit the original contributors of research data receive if their data is used. Contributions from developed countries to various other countries, such as developing and underdeveloped countries, can share and reuse research data, significantly benefiting the countries.

Conclusion

Research is the central role of all subjects of knowledge, and it entails a variety of intellectual tasks that are both time-consuming and cost-effective. The study's findings concluded that the RDM trend in the research environment is on the rising trend. The research project's product is research publications, and data gathered in any format using various data collecting tools for the research findings are vanishing by the study itself. There is no mechanism for keeping such data in a single source, notably in India and other developing nations, which used to invest enormous amounts of money, time, and resources in research but have a poor record of maintaining the data. Other developed countries are preserving and reusing research data, while some developing countries, such as China, Spain, Africa, and Iran, are initiating and practicing preserving and reusing research data, which is a valuable source in the Research Life Cycle. Therefore, it is suggested that developing countries like India adopt the strategies the developed countries followed to preserve the research data and make it visible and reused. This will enhance research transparency and lead to higher productivity in every field.

References

- Alonso, J. M., Castiello, C., & Mencar, C. (2018, June). A bibliometric analysis of the explainable artificial intelligence research field. In *International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems* (pp. 3-15). Springer, Cham. https://doi.org/10.1007/978-3-319-91473-2_1
- Anilkumar, N. (2018). Research data management in India: a pilot study. In *EPJ Web of Conferences* (Vol. 186, p. 03002). EDP Sciences. <https://doi.org/10.1051/epjconf/201818603002>

- Ashiq, M., Usmani, M. H. & Naeem, M. (2020). A systematic literature review on research data management practices and services. *Global Knowledge, Memory, and Communication*. <https://doi.org/10.1108/GKMC-07-2020-0103>
- Bangani, S. & Moyo, M. (2019). Data sharing practices among researchers at South African universities. *Data Science Journal*, 18(1). <https://doi.org/10.5334/dsj-2019-028>
- Biancone, P. Pietro, Saiti, B., Petricean, D. & Chmet, F. (2020). The bibliometric analysis of Islamic banking and finance. *Journal of Islamic Accounting and Business Research*, 11(10), 2069-2086. <https://doi.org/10.1108/JIABR-08-2020-0235>
- Bibliometrix R Package. (n.d.). Retrieved December 23, 2020, from <https://www.bibliometrix.org/index.html>
- Bunkar, A. R. & Bhatt, D. D. (2020). Perception of researchers & academicians of parul university towards research data management system & role of library: A study. *Desidoc Journal Of Library & Information Technology*, 40(3), 139–146. <https://doi.org/10.14429/djlit.40.3.15302>
- Chawinga, W. D. & Zinn, S. (2020a). Research data management at an African medical university: Implications for academic librarianship. *Journal Of Academic Librarianship*, 46(4). <https://doi.org/10.1016/j.acalib.2020.102161>
- Chawinga, W. D. & Zinn, S. (2020b). Research data management in universities: A comparative study from the perspectives of librarians and management. *International Information and Library Review*, 53(2), 97-111. <https://doi.org/10.1080/10572317.2020.1793448>
- Chiware, E. R. T. (2020). Data librarianship in South African academic and research libraries: a survey. *Library Management*, 41(6-7), 401-416. <https://doi.org/10.1108/LM-03-2020-0045>
- Cho, J. (2019). Study of Asian RDR based on re3data. *Electronic Library*, 37(2), 302-313. <https://doi.org/10.1108/EL-01-2019-0016>
- Choi, M. S. & Lee, S. (2020). Research data management status of science and technology research institutes in Korea. *Data Science Journal*, 19(29), 1-11. <https://doi.org/10.5334/DSJ-2020-029>
- Corrall, S., Kennan, M. A. & Afzal, W. (2013). Bibliometrics and research data management services: Emerging trends in library support for research. *Library Trends*, 61(3), 636-674. <https://doi.org/10.1353/lib.2013.0005>
- Cox, A. M. & Tam, W. W. T. (2018). A critical analysis of lifecycle models of the research process and research data management. *Aslib Journal Of Information Management*, 70(2), 142–157. <https://doi.org/10.1108/AJIM-11-2017-0251>
- Cyberinfrastructure Council (2007). *Cyberinfrastructure vision for 21st century discovery*. National Science Foundation. Retrieved from <http://www.nsf.gov/pubs/2007/nsf0728/nsf0728.pdf>
- Fusco, F., Marsilio, M. & Guglielmetti, C. (2020). Co-production in health policy and management: A comprehensive bibliometric review. *BMC Health Services Research*, 20(1), 504. <https://doi.org/10.1186/s12913-020-05241-2>
- Gunjal, B. & Gaitanou, P. (2017). Research data management: A proposed framework to boost research in higher educational institutes. *IASSIST Quarterly*, 41(1–4), 12. <https://doi.org/10.29173/iq12>

- Herbst, K., Juvekar, S., Bhattacharjee, T., Bangha, M., Patharia, N., Tei, T., Gilbert, B. & Sankoh, O. (2015). The INDEPTH Data Repository: An International Resource for Longitudinal Population and Health Data From Health and Demographic Surveillance Systems. *Journal of Empirical Research on Human Research Ethics (JERHRE)*, 10(3), 324–333. <https://doi.org/10.1177/1556264615594600>
- Janik, A., Ryzko, A. & Szafraniec, M. (2020). The scientific landscape of smart and sustainable cities literature: A bibliometric analysis. *Sustainability*, 12(3), 779. <https://doi.org/10.3390/su12030779>
- Joo, S. & Peters, C. (2020). User needs assessment for research data services in a research university. *Journal of Librarianship and Information Science*, 52(3), 633–646. <https://doi.org/10.1177/0961000619856073>
- Koltay, T. (2017). Research 2.0 and research data services in academic and research libraries: priority issues. *Library Management*, 38(6–7, S.I.), 345–353. <https://doi.org/10.1108/LM-11-2016-0082>
- Kun, W., Tong, L. & Xiaodan, X. (2019). Application of big data technology in scientific research data management of military enterprises. *Procedia Computer Science*, 147, 556–561. <https://doi.org/10.1016/j.procs.2019.01.221>
- Late, E. & Kekalainen, J. (2020). Use and users of a social science research data archive. *PLOS ONE*, 15(8), e0233455. <https://doi.org/10.1371/journal.pone.0233455>
- Liao, Y. P. & Ma, T. J. (2018). Institutional repositories: A bibliometric study of the social sciences citation index (SSCI). *Electronic Library*, 36(3), 504–517. <https://doi.org/10.1108/EL-09-2017-0192>
- Liu, G., Zotoo, I. K. & Su, W. (2020). Research data management policies in USA, U.K. and Australia universities: An online survey. *Malaysian Journal of Library & Information Science*, 25(2), 21–42. <https://doi.org/10.22452/mjlis.vol25no2.2>
- Liu, H., Peng, K., Li, W. & Cao, Y. (2019). Investigation on the trends and characteristics of articles on submerged macrophytes: perception from bibliometrics between 1991 and 2018. *Journal of Freshwater Ecology*, 34(1), 703–713. <https://doi.org/10.1080/02705060.2019.1676319>
- Liu, X. & Ding, N. (2016). Research data management in universities of central China: Practices at Wuhan University Library. *Electronic Library*, 34(5), 808–822. <https://doi.org/10.1108/EL-04-2015-0063>
- Marlina, E. & Purwandari, B. (2019). Strategy for research data management services in Indonesia. *Procedia Computer Science*, 161, 788–796. <https://doi.org/10.1016/j.procs.2019.11.184>
- Mushi, G. E., Pienaar, H. & Deventer, M. van. (2020). Identifying and implementing relevant research data management services for the library at the University of Dodoma, Tanzania. *Data Science Journal*, 19(1), 1–9. <https://doi.org/10.5334/dsj-2020-001>
- Onyancha, O. B. (2016). Open research data in Sub-Saharan Africa: A bibliometric study using the data citation index. *Publishing Research Quarterly*, 32(3), 227–246. <https://doi.org/10.1007/s12109-016-9463-6>
- Payal, Awasthi, S. & Tripathi, M. (2019). A selective review of literature on research data management in academic libraries. *DESIDOC Journal of Library and Information Technology*, 39(6), 338–345. <https://doi.org/10.14429/djlit.39.06.14451>

- Penn State University Libraries (2014). *What is data management?* Pennsylvania State University Libraries, State College, PA. Retrieved from https://web.archive.org/web/20160304192954/http://www.libraries.psu.edu/psul/pubcur/what_is_dm.html
- R Core Team (2018). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. Retrieved from <https://www.R-project.org/>.
- Redkina, N. S. (2019). Current trends in research data management. *Scientific And Technical Information Processing*, 46(2), 53-58. <https://doi.org/10.3103/S0147688219020035>
- Roche, D. G., Granados, M., Austin, C. C., Wilson, S., Mitchell, G. M., Smith, P. A., Cooke, S. J. & Bennett, J. R. (2020). Open government data and environmental science: A federal Canadian perspective. *Facets*, 5(1), 942–962. <https://doi.org/10.1139/FACETS-2020-0008>
- Ruiz-Rosero, J., Ramirez-Gonzalez, G., Williams, J. M., Liu, H., Khanna, R. & Pisharody, G. (2017). Internet of things: A scientometric review. *Symmetry*, 9(12), 1-34. <https://doi.org/10.3390/sym9120301>
- Saeed, S. & Ali, P. M. N. (2019). Research data management and data sharing among research scholars of life sciences and social sciences. *DESIDOC Journal Of Library & Information Technology*, 39(6), 290-299. <https://doi.org/10.14429/djlit.39.6.14997>
- Secinaro, S., Brescia, V., Calandra, D. & Biancone, P. (2020). Employing bibliometric analysis to identify suitable business models for electric cars. *Journal of Cleaner Production*, 264, 121503. <https://doi.org/10.1016/j.jclepro.2020.121503>
- Surkis, A. & Read, K. (2015). Research data management. *Journal of the Medical Library Association*, 103(3), 154-156. <https://doi.org/10.3163/1536-5050.103.3.011>
- Tripathi, M., Chand, M., Sonkar, S. K. & Jeevan, V. K. J. (2017). A brief assessment of researchers' perceptions towards research data in India. *IFLA journal*, 43(1), 22-39. <https://doi.org/10.1177%2F0340035216686984>
- Tripathi, M., Shukla, A. & Sonker, S. K. (2017). Research data management practices in university libraries: A study. *DESIDOC Journal of Library and Information Technology*, 37(6), 417-424. <https://doi.org/10.14429/djlit.37.6.11336>