Growth and development of Research Data Management in India: A visual analysis of published literature indexed in Scopus

Manu T R
1Assistant Librarian, Central Library, Indian Institute of Technology Delhi, Hauz Khas, New Delhi, Delhi 110016 & Research Scholar, School of Library and Information Science, Central University of Gujarat, Sector-29, Gandhinagar – 382030, manutr91@gmail.com

Bhakti Gala
2Assistant Professor, School of Library and Information Science, Central University of Gujarat, Sector-29, Gandhinagar – 382030., bhakti.gala@cug.ac.in

Growth and development of Research Data Management in India: A visual analysis of published literature indexed in Scopus

Manu T R[1*] and Bhakti Gala[2]

[1]Assistant Librarian, Central Library, Indian Institute of Technology Delhi, Hauz Khas, New Delhi, Delhi 110016 & Research Scholar, School of Library and Information Science, Central University of Gujarat, Sector-29, Gandhinagar – 382030

[2]Assistant Professor, School of Library and Information Science, Central University of Gujarat, Sector-29, Gandhinagar – 382030.

[*]Email: manutr91@gmail.com

Abstract

The present paper identifies and presents a visual analysis of Research Data Management (RDM) literature published by Indian researchers and indexed in Scopus – the bibliographic and citation database. The researchers have used standard bibliometric methods of measurements to conduct the study such as Bradford's law, Lotka's law and Zipf's Law among other parameters. The study findings indicate that India is rapidly focusing on RDM with its contribution of 136 publications to the global output of 4929 publications over a period of 1945 to 2021, ranking 12th globally. Out of 136 publications, 35.29% (48) were published as open access publications and the remaining 30.15%, (41) were published as gold, 25.74% (35) were green and 4.41% (6) were published as hybrid gold and bronze access respectively. Research data management, big data, data curation, data repositories, digital curation, research data sharing, data privacy, data security, data stewardship, data preservation are major research trends discussed within the broad area of RDM., This study highlights the contribution of LIS professionals to RDM research, although RDM research is still at a nascent stage in India. This study charts the growth of published RDM literature in India since 2015. Researchers irrespective of their disciplines are acquiring the necessary skills and expertise to prepare themselves to embrace this new area.

Keywords: Research Data Management, Data curation, Bibliometric Study, Scopus, Libraries, Visual analysis, India

1 Introduction

Research data management (RDM) is at the core of successful research and is globally perceived as part of the good research practice of researchers. In this data-driven society, RDM is the process that guides researchers through the different stages of the research data lifecycle including planning, collecting, analyzing, publishing, preserving, sharing, and reusing (Petrovich, 2020). Research data is generated from original research results, publication research, theses and dissertation, and other research processes. Evidence of the importance of RDM can be seen in the increasing number of data journals, data citations that describe datasets (Surkis and Read, 2015). RDM deals with data collection methods, best practices and implementations of infrastructure and services to archive preserve and make research data re-usable. It addresses the full range of information needs including data storage, data security, preservation, compliance, quality of data and data sharing. However, researchers have their concerns, practices, perceptions, and attitudes on managing and sharing research data, and they may need convincing before they are willing to share their research data with institutional repositories (Manu, 2018). Therefore, it is necessary for institutes and universities with the collaboration of researchers and libraries to understand the researcher's perception, level of researcher's awareness on RDM and develop institutional data repositories and data sharing policy guidelines that accept the research data generated at the institutional level.

RDM supports an open access ecosystem to share research data through open access data repositories especially research funded from public funds. Thus, recently Indian research funding agencies from the Government of India like Department of Science and Technology, Indian Council of Agricultural Research (ICAR), Indian Council of Social Science Research (ICSSR), Department of Biotechnology, and International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), have formulated data management and sharing policies for research data generated from government funded research to mandate proper data management throughout the research lifecycle stating guidelines for sharing of research data through research data repositories to make this data publicly accessible.

Bibliometrics as a method for measuring, monitoring, and studying scientific output is well established (Oliveira, 2019). Pritchard (1969) first introduced the word Bibliometrics and defined it as, "the application of mathematical and statistical methods to books and other media of communication". It is an imperative scientific tool to evaluate the impact of scientific output using quantitative measures. Further, it involves the measurement of the physical unit of publications, citation count and H-index, etc. Bibliometric studies are conducted by applying

three basic laws namely: Lotka's Law (Productivity of authors in terms of scientific papers), Bradford's Law (Scattering of articles over different journals), and Zipf's Law (Frequency of occurrence of words in the text).

2   Literature Review

Bibliometric research has been conducted in various disciplines to identify Indian output in the subject of crystallography (Selvi and Gopalakrishnan, 2016), LIS research (Bhakta, Kar and Bhui, 2019), drug abuse (Azmi and Abbas, 2018), social science and humanities (Tripathi, Kumar and Babbar, 2018), and ethnobotany (Pathak and Bharat, 2020). These studies analyzed literature drawn from specific bibliographic databases such as Scopus (Selvi and Gopalakrishnan, 2016; Azmi and Abbas, 2018; Pathak and Bharat, 2020) and indexes such as Indian Citation Index Database (Bhakta, Kar and Bhui, 2019), Social Science Citation Index (SSCI), and Arts & Humanities Citation Index (A&HCI) (Tripathi, Kumar and Babbar, 2018). These studies analyzed the retrieved literature using various bibliographic parameters such as publication growth, year-wise publications, document types, authorship patterns, research trends, core journals, subject wise analysis, citation analysis, geographical contribution etc. (Selvi and Gopalakrishnan, 2016; Bhakta; Kar and Bhui, 2019; Azmi and Abbas, 2018; Tripathi, Kumar and Babbar, 2018; Pathak and Bharat, 2020).

Bibliometric studies on global scholarly contribution on RDM (Zhang and Eichmann-Kalwara, 2019) and Sub-Saharan African (SSA) researcher's contribution on research data sharing (Onyancha, 2016) were mapped according to the literature indexed by Scopus (Zhang and Eichmann-Kalwara, 2019) and Data Citation Index (Onyancha, 2016). The study (Zhang and Eichmann-Kalwara, 2019) further found that 1913 documents containing 23402 cited references were published from 1962 to 2018 on a wide range of subjects in various formats. In addition, by sharing their research data 61.3% of research contributors in the SSA region were able to achieve high impact in their research output, an improvement of scientific transparency and accuracy, increased research collaboration and an increase in research accessibility (Onyancha, 2016).

In India, RDM is an emerging area that has drawn the attention of researchers in higher education institutes with RDM research being at an initial stage (Bhardwaj, 2019; Tripathi and Sonker, 2017). Several studies have been conducted to understand its impact in state-of-the-art scholarly research and its role in research data organizing, data curation, preserving raw data for long-term and creating awareness of research activities (Tripathi and Sonker, 2017; Anilkumar,

2018; Saeed and Ali, 2019). The selected global literature review was undertaken on RDM in academic libraries that focused on concepts, policies, data caution, data sharing, legal and ethical issues, copyright issues, the role of libraries and availability of RDM services in libraries (Payal and Tripathi, 2019; Tripathi and Sonker, 2017; Singh, Monu, and Dhingra, 2018). The conceptual framework of RDM for higher educational institutes (Tripathi and Pandy, 2018; Singh, Monu, and Dhingra, 2018) and the national repository of open research data (Patel, 2016) present the workflow of research data life cycle in its various phases. They compared the openly available 14 DMP tools with parameters to assist researchers and data managers to formulate effective data management plans (Gajbe et al., 2021). A bibliometric and scientometric study was carried out on RDM literature at global level (Gupta, Arora, Chakravarty, 2021). Citations and visibility are key factors that motivate researchers to share data (Saeed and Ali, 2019). Research funding agencies in India such as DST Government of India (DST, 2012), ICRISAT (ICRISAT, 2014) and ICAR (ICAR, 2014) have come out with policies on data management, with a focus that public-funded research should be made available for public access without restriction, therefore researchers are required to share data before or immediately after completion of research projects.

The above literature review illustrates bibliometric studies conducted to analyze the impact of various research outputs using quantitative measures. Although several studies have been conducted by international scholars, not a single study has been done to identify published RDM literature from India. This is the first study undertaken to identify the contribution of Indian researchers to RDM published literature, indexed in international bibliographic and citation databases. This study will benefit researchers, policymakers, and government funding agencies to be able to get an overview of the disciplines in which RDM is developing at the national stage. This study also aims to help collaborators from the developing world to become aware of the development and growth of RDM publications in India. After two decades since the advent of RDM, such a study is required to chart out the growth and development of RDM literature in India.

3 Objectives

The objectives of this research study are governed by the primary research question: What is the contribution of Indian researchers in Research Data Management? The broad objectives of the study are as follows:

- To highlight the contribution of Indian researchers towards RDM publications.

- To identify the annual literature growth, prolific authors, subject coverage, top-cited research papers, frequently published sources, and institute contributions on RDM.

- To display international collaborative research in RDM.

- To understand the emerging research areas explored in the literature published by Indian research scholars relating to managing research data.

- To identify the LIS professional contribution on RDM in India

4   Research Methodology

The researchers adopted quantitative bibliometrics research methods to carry out this study. A quantitative bibliometric study is used to map scholarly literature to reveal research impact and trends. It is based on a sufficiently large sample size with literature data exceeding at least 50 articles (Sjöstedt, Aldberg, and Jacobsson, 2015). The researchers used the bibliographic database Scopus to retrieve published literature on RDM by Indian researchers. The selection of keywords was intended to perform more focused searches on the entire RDM life cycle. Therefore, researchers used the different stages of the RDM life cycle *i.e.,* research planning, data collection, data analysis, publishing/sharing, preserving, and reusing as keywords for searching the database. The search was performed using advanced search filters available in Scopus. To filter Indian publications, the filter AFFILCOUNTRY ("India") was used. Keywords given by the authors were further counted to identify research trends in RDM. The published works of literature on RDM spanning a period of 28 years from January 1993 to December 2021 were retrieved for the study as of 31st October 2021.

The VOSviewer open-source software tool for analyzing and visualizing bibliometric networks (Centre for Science and Technology Studies, 2020) and advanced excel graphics was used for presenting the visual analysis of data retrieved from the Scopus database. The VOSviewer was used for author keywords and the geographical collaboration network of Indian researchers in RDM published literature. Author keywords were extracted from the bibliographic information available on the Scopus database and these were further analyzed for co-occurrence by using VOSviewer to identify the major trends in RDM research. Scopus analysis graphics were used for H-Index graphics, top ranked profile authors, subject area, and documents per year by sources. The data retrieved was also analyzed using different bibliometric parameters including

laws such as Lotka's Law (prolific authors, subject wise contribution, institute wise contribution and geographical collaborations), Bradford's Law (growth of published literature, cited literature and a top source of Literature), Zipf's Law (author keywords analysis) and tables and graphics were created using advanced excel graphics.

5   Limitations of the study

The current research study is limited to RDM published literature by Indian scholars and indexed by the Scopus bibliographic citation database. Scopus was selected as a preliminary analysis to show that it indexed more published literature on RDM by Indian scholars instead of the other leading citation database *i.e.,* Web of Science. The study is also limited to results retrieved using two fields available in the advanced search options available on the Scopus database namely TITLE-ABS-KEY and AFFILCOUNTRY. A combination of different keywords was used for searching the TITLE-ABS-KEY fields viz. TITLE-ABS-KEY ("research data manag*") OR TITLE-ABS-KEY ("responsible data manag*") OR TITLE-ABS-KEY ("research support data*") OR TITLE-ABS-KEY ("data lifecycle manag*") OR TITLE-ABS-KEY ("data resource manag*") OR TITLE-ABS-KEY ("research data admin*") OR TITLE-ABS-KEY ("digital curat*") OR TITLE-ABS-KEY ("digital data manag*") OR TITLE-ABS-KEY ("data steward*") OR TITLE-ABS-KEY ("data curat*") OR TITLE-ABS-KEY ("research reposit*") OR   TITLE-ABS-KEY ("data management plan*") OR TITLE-ABS-KEY ("RDM in libraries*") OR   TITLE-ABS-KEY ("research data literacy*") OR   TITLE-ABS-KEY ("DMPTookl*") OR TITLE-ABS-KEY ("research data sharing*") OR TITLE-ABS-KEY ("research data preservation*") OR   TITLE-ABS-KEY ("data management policy*") OR TITLE-ABS-KEY ("RDM services*") OR TITLE-ABS-KEY ("FAIR principles*") OR TITLE-ABS-KEY ("data citation*") AND AFFILCOUNTRY ("India").

6   Data Analysis and Findings

The research findings of the present study demonstrate that RDM and its services are demonstrating definite growth. Researchers, practitioners, decision-makers, funders, and library professionals have started exploring the value of research data, the changing role of libraries, different services to support research scholars and the requirements of faculty members for training and education. The complete published literature on RDM found around 4929 publications from 114 countries around the world, where the United States stands at 1st position

with 1935 (39.26%) literature and India stands at 12th position with 136 (2.76%) publications (Figure 1). It indicates that interest in RDM in India is growing rapidly.

Figure 1. Global geographical contributions on RDM Research

6.1.Literature on Research Data Management in India

The earliest RDM related published literature written by an Indian researcher on oil and natural gas research data was published in January 1993. Table 2 presents a summary of published literature found in Scopus on RDM written by Indian scholars. Total 136 articles were published by Indian scholars, which have received 1561 citations until 31st October 2021, at the average citation of 11.47 per article with an h-*index* of 17 presented as Scopus h-graph (Figure 2).

Table 2. RDM Literature Published by Indian researchers

Total no. of literature	136
Total citations received	1561
Average citation per article	11.47
h-index	17

Figure 2. Scopus h-graph


These documents h-index17

Of the documents considered for the h-index, have been cited at least times
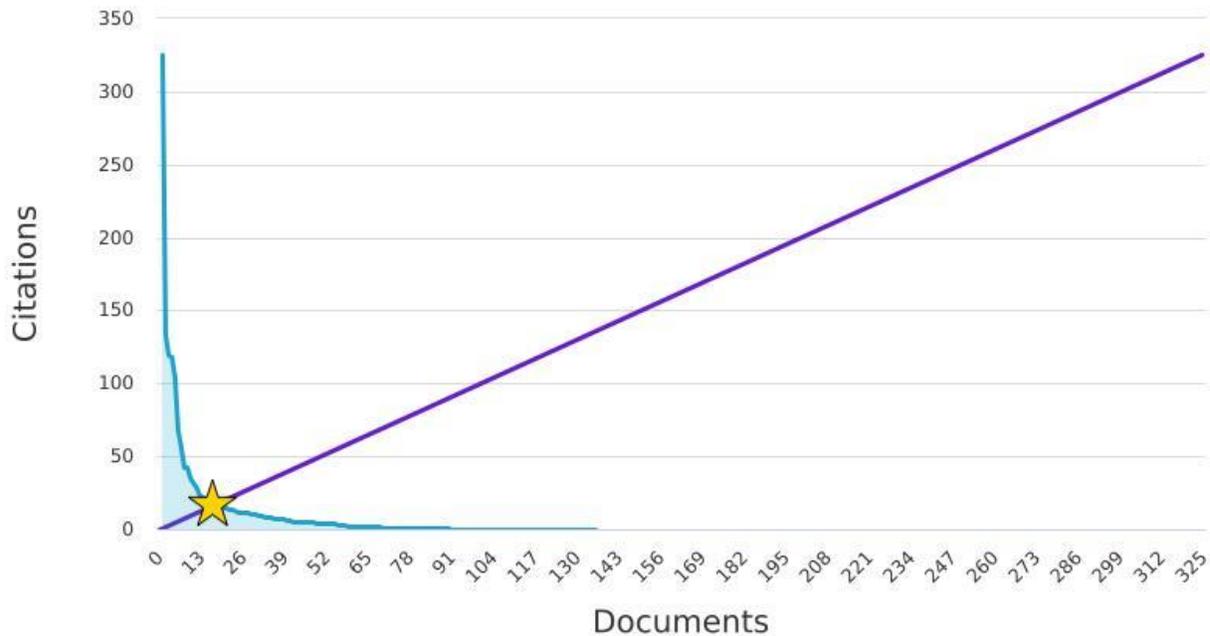
6.2. Annual Growth of Literature and Citations

The bibliometric parameter of annual growth studies, and Figure 3 displays the year-wise literature, citation growth and the cumulative total of published literature on RDM by Indian researchers. Of the total 136 works of literature until now, 35.29% (48) were published as open access literature, 30.15%, (41) were published as gold, 25.74% (35) were green and 4.41% (6) were published as hybrid gold and bronze access respectively. These have been distributed in different document types such as journal (100 literature), conference papers (27 literature), book series (6 literature), two trade journals and one book. The literature related to RDM had been publishing since 1993, but no significant growth was observed until 2015. A steady increase in the literature since 2015 and average per year publication increased 185.71% articles during 2016 to 2017, followed by 123.07% and 162.5% publication have increased in the year 2017 to 2018 and 2018 to 2019 respectively. Total 1561 citations were received and maximum of 374 citations were received by this literature in 2021 followed by 319 citations in 2020, 232 citations in 2019, 191 citations in 2018, 118 citations in 2017 and 130 citations were received in 2016. It demonstrates that the impact of RDM published literature by Indian scholars is gradually influencing more citations. The growth of literature is evidence of the growing interests of Indian researchers in exploring and adopting RDM practices.

**Figure 3. Annual growth literature and citations**
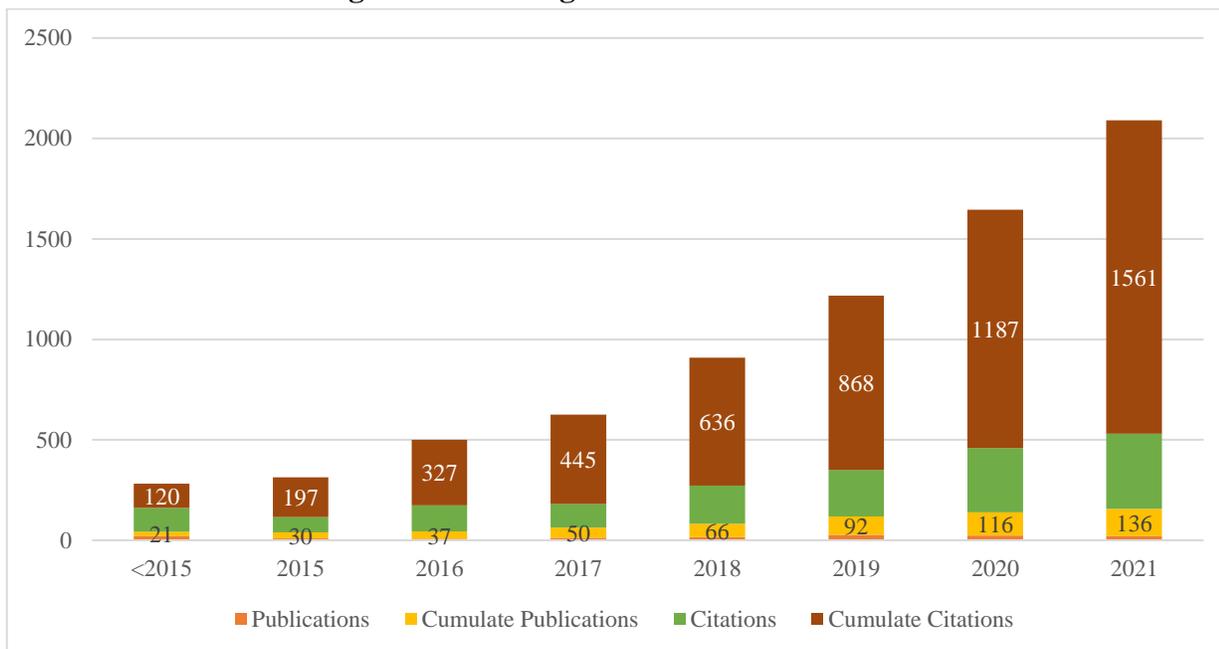
6.3.Prolific Indian Authors Contribution to RDM Literature

Identifying the top prolific authors in RDM published literature would help in identifying the Indian researchers'/authors'/practitioners' active contribution to RDM research in India. A total of 845 authors have contributed to RDM research in India from around the world. The most prolific authors (Figure 4) are from the disciplines of computer science and engineering, cancer care, education, and machine intelligence. Accordingly, authors from engineering and technological institutes have more contributions. Gadad, V. and Sujatha, B have individually published a maximum of 03 (2.20%) literature. These authors represent the engineering colleges: R.V College of Engineering, Bengaluru, Karnataka and Sengunthar Engineering College, Tiruchengode, Tamilnadu respectively. There are 30 individual authors who have contributed 2 (1.47%) articles each. These authors are from Jadavpur University, Kolkata; Orchids Breast Health Clinic, Pune; Periyar University, Salem; Indian Statistical Institute, Kolkata and the University of Delhi etc.

Figure 4. Top ranked prolific Indian authors in RDM published literature

Documents by author

Compare the document counts for up to 15 authors.
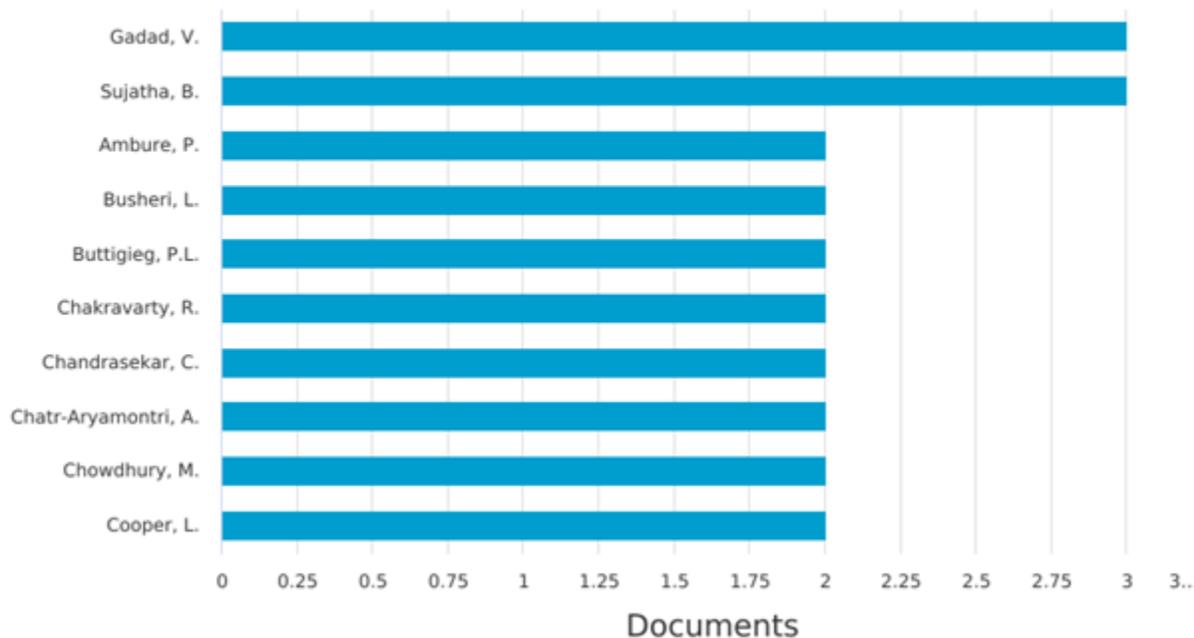
6.4. Subject-Scope of Literature

The analysis of the data illustrates that published literature on RDM by Indian researchers is multidisciplinary. Researchers and practitioners from various subjects have been involved in exploring RDM services, its impact, and its value on their research. The Scopus subject classification was used to classify the literature into major subject groups. Scopus has classified this literature under 16 broad subjects, majority of RDM literature (Figure 5) *i.e.,* 65 (47.79%) is from computer science followed by social science with 33 (24.26%) literature. The 23 (16.91%) and 20 (14.71%) literature published related to biochemistry, genetics and molecular biology and medicine discipline respectively. Over 31 (14.9%) literature are covered in the other subject category as per Scopus classification that includes *physics and astronomy; business, management and accounting; multidisciplinary; neuroscience; chemistry; earth and planetary sciences; energy environmental science; chemical engineering; health professions; immunology and microbiology; materials science; psychology and veterinary*. Analysis of the subject-wise RDM literature indicates that computer science is a major discipline where RDM has been explored in-terms of data repository development, metadata schema, data curation, data archive and preservation etc.

Figure 5. Subject-area wise distribution of RDM literature

Documents by subject area
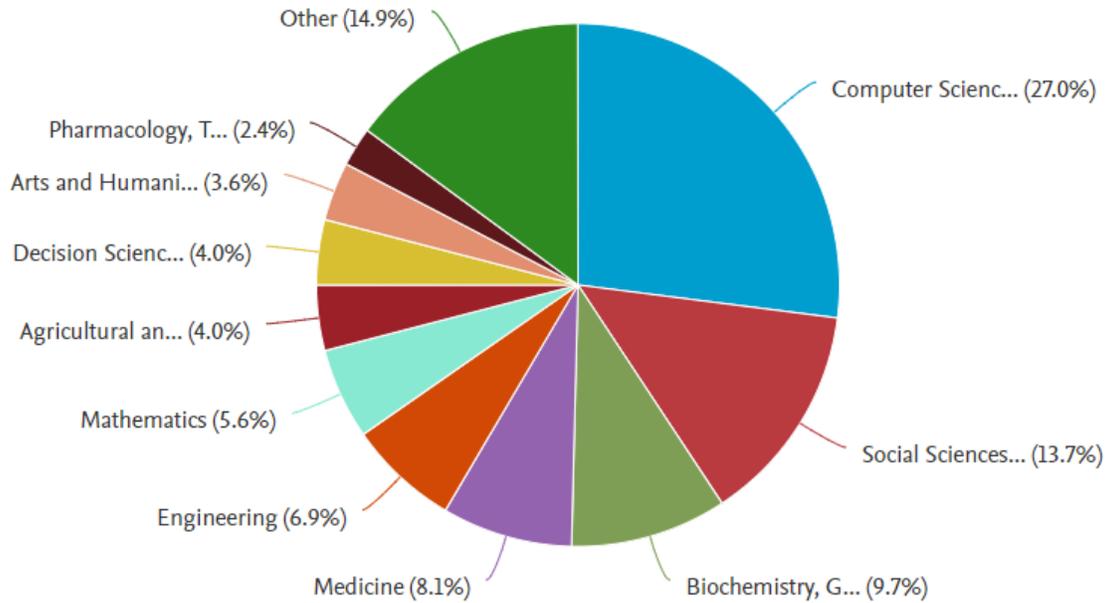
6.5.Top Cited Literature

Citations are evidence of literature that has been making an impact on another research. Total 1561 citations have been received by the RDM published literature by Indian scholars at an average of 11.47% per article. These citations have been received from 90 literatures (66.17%) and the remaining 46 works of literature (33.82%) are yet to receive citations. The top cited RDM published literature with over 100 citations (Table 1) are from the scientific peer reviewed journals such as Nature methods (20.82% citations), PLoS biology (8.52% citations), Molecular informatics (7.62% citations), Nucleic acids research (7.55% citations) and Software: practice and experience (6.66% citations). These top-cited works of literature focused on data curation, understanding of different landscape types of data, data analysis and development of data repositories. It is interesting that although most works of literature are from computer science (Fig. 5), only one publication in the top 5 is from computer science.

Table 1. Top cited literature

| Authors	Title	Source	Vol. issue no. and the page no.	Year	Publisher	No. of citations
Orchard, S., et al.	Protein interaction data curation: The International Molecular Exchange (IMEx) consortium	Nature Methods	9, (4) and 345-350	2012	Nature	325 (20.82%)
Deans, A. R., et al.	Finding Our Way through Phenotypes.	PLoS Biology	13(1)	2015	PLoS	133 (8.52%)
Gramatica, P., et al	QSAR Modeling is not "Push a Button and Find a Correlation": A Case Study of Toxicity of (Benzo-) triazoles on Algae	Molecular Informatics	31, (11–12) and 817-835	2012	Wiley Online Library	119 (7.62%)
Urban, M., et al.	PHI-base: A new interface and further additions for the multi-species pathogen–host interactions database	Nucleic Acids Research	45, (D1, D604-D610	2017	Oxford University Press	118 (7.55%)
Kune, R., et al.	The anatomy of big data computing: Anatomy of Big Data Computing	Software: Practice and Experience	46, (1) 79-105	2016	Wiley Online Library	104 (6.662%)

6.6. Top Sources of Literature

The top sources of RDM published literature are from various formats including peer-reviewed journals, lecture notes, book chapters, conference papers, reviews etc. RDM literature authored by Indian researchers is found in the 83 different types of sources of literature. Figure 6 depicts the top 06 sources of publication that have published more than three works of literature on RDM. The major sources such as Database: The Journal of Biological Databases and Curation by OUP have published a maximum of 8 (5.88%) literature and Library Philosophy and Practice (LPP) and digital commons have published 6 (4.41%) literature. *DESIDOC Journal of Library and Information Technology* from DRDO has published the 5 (3.68%), ACM International Conference Proceeding Series, Nucleic Acids Research an OUP journal and Scientific Reports by Nature have published 3 (2.21%) literature individually. Major publication sources like Cluster Computing (Springer journal), Global Knowledge, Memory and Communication, Indian Journal of Public Health Research and Development, International Journal of Soft Computing, Lecture Notes in Computer Science, Library Hi-Tech, Library Philosophy and Practice and

Music Reference Services Quarterly also have equal contributions of 2 publications each. It also illustrates that maximum publications are from journals (73.53%) followed by conference proceedings (19.85%).

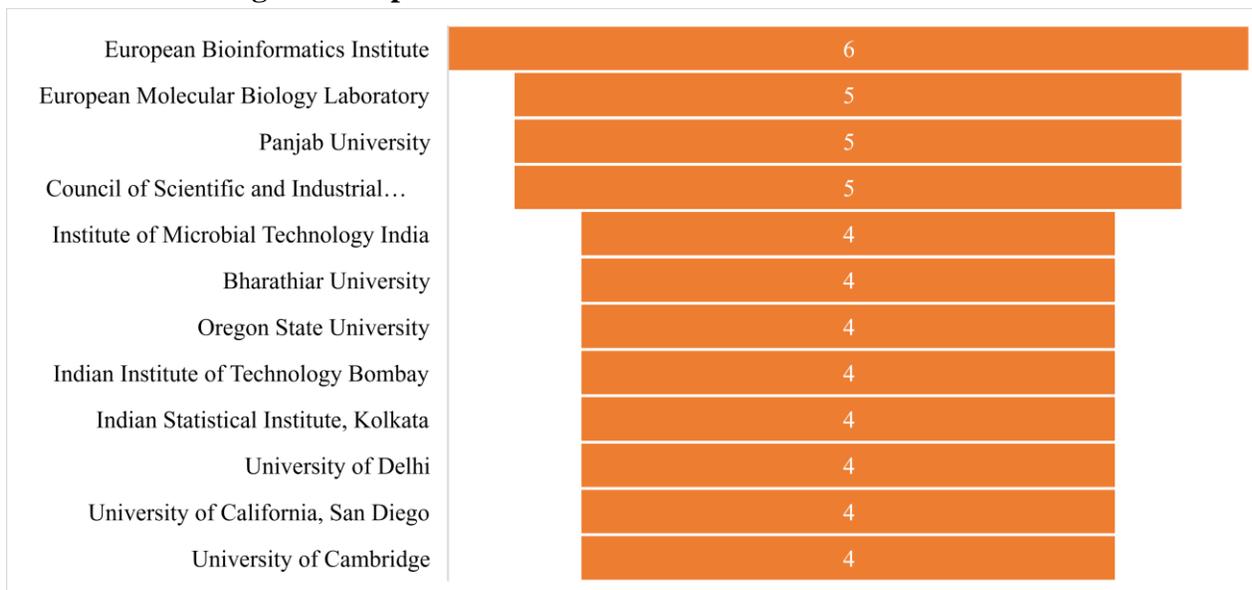**Figure 6. Top sources of published RDM literature**
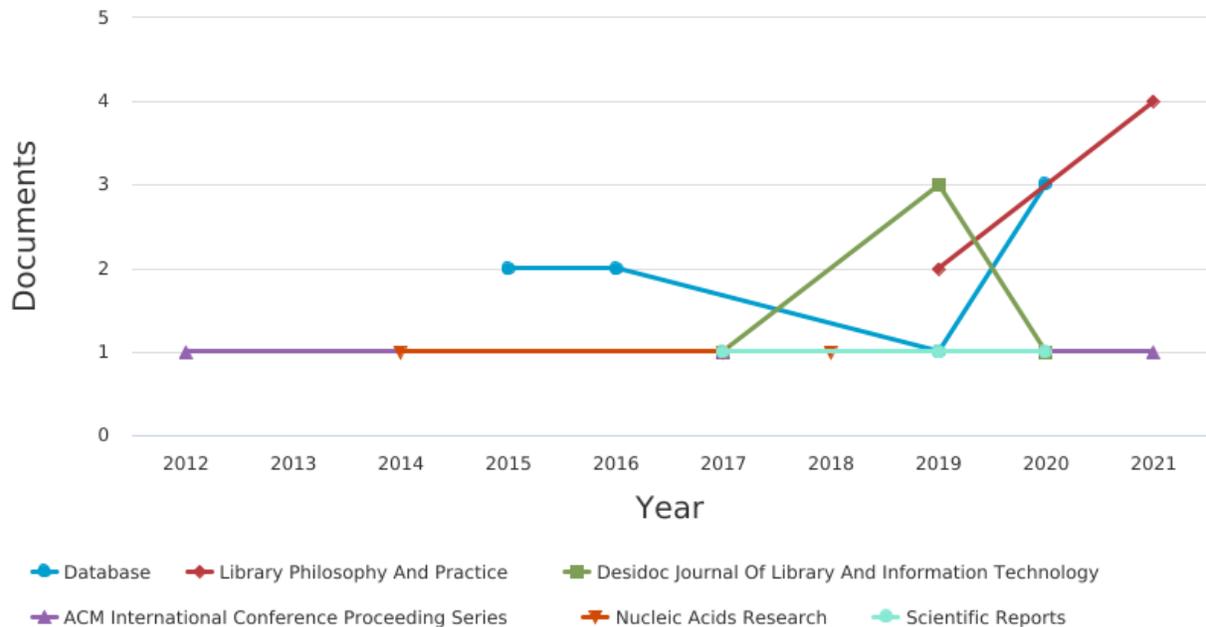
6.7. Institute-Wise Contribution of RDM Literature

Academic and research institutions in India have steadily been involved in exploring the management of research data, tools, and technologies available to develop RDM services and data repositories at the institutional level. The study has found that researchers from 160 affiliated academic and research institutions have contributed to RDM published literature in India. Of the total 160 institutions contributing to RDM publications, 51 (31.87 %) are research institutes and the remaining 109 (68.12%) are academic institutions. Figure 7 depicts that major contributing institutes include European Bioinformatics Institute with a maximum contribution of 6 (4.41%) followed by the European Molecular Biology Laboratory, Panjab University, Council of Scientific and Industrial Research India equally contributing the 5 (3.68%) works of literature each. Institute of Microbial Technology India, Bharathiar University, Oregon State University, Indian Institute of Technology Bombay, Indian Statistical Institute, Kolkata, University of Delhi, University of California, San Diego and the University of Cambridge have contributed with 4 (2.94%) literature each. Research institutes or science laboratories like

Orchids Breast Health Clinic, Prashanti Cancer Care Mission have a maximum contribution of 2 (1.85%) RDM related publications.

Figure 7. Top Institute wise contribution of RDM literature

| Institute	Contribution
European Bioinformatics Institute	6
European Molecular Biology Laboratory	5
Panjab University	5
Council of Scientific and Industrial…	5
Institute of Microbial Technology India	4
Bharathiar University	4
Oregon State University	4
Indian Institute of Technology Bombay	4
Indian Statistical Institute, Kolkata	4
University of Delhi	4
University of California, San Diego	4
University of Cambridge	4

6.8. Geographical Collaborations and Network

Collaboration plays an important role in successful interdisciplinary or multidisciplinary research outcomes. Scientific collaboration and network partnership are key components of the institutes' research activities to produce impactful research. The geographical (countries) collaborations and network with Indian researchers are presented in Figure 7. The maximum collaboration in RDM published literature is with the United States (USA) with 24 (17.64%) followed by the United Kingdom (UK) with 18 (13.23%) and Germany with 12 (8.82%) literature. Indian researchers have collaborated with over 50 countries across the world in publishing RDM published literature. The data findings also indicate that other countries like France, Italy, Australia, Canada, Netherlands, Norway, Switzerland, Greece, Malaysia and Sweden (more than 4 articles) are also actively contributing their research projects to RDM.

Figure 7. Geographical Collaborations and Network of RDM literature in India

6.9.Major Areas of RDM Research in India

 The present study suggests that research on RDM is in the early stages in India. The published literature examined by this bibliometric study indicates that 84.55% (105 literatures) of articles have been published after 2015. Since then, the exponential growth in the published literature reflects the potential of an increase in the importance of RDM in India. An analysis of the keyword's occurrences helped the researchers to identify the core research areas of RDM studied in the published literature. Total over 2213 keywords are found, of which 1345 (60.77%) are unique keywords. As presented in figure 8, 'research data management' including RDM services occurs in 18 (13.24%) RDM published literature followed by the 'big data' that also appeared in 34 (25.00%) publications, and 'digital curation', 'data repositories', 'data curation' and 'research data' is found in 23 (16.91%) publications equally. Other RDM related topics are data privacy, data security, data stewardship, data preservation, data sharing, data type and data analysis. They have also been significantly discussed in the RDM published literature by Indian scholars. The subject scope of the RDM published literature includes research on data sharing among faculty members, researcher perception, data generation, data access, the role of libraries in providing innovative data management services including the development of policy guidelines, data

curation and data repositories. It is also found that in the RDM published literature an attempt is made to describe institutional RDM policy, institutional framework, roles, and responsibilities of RDM stakeholders including management, research support centers, research supervisors, researchers, and libraries.

**Figure 8. Author keyword occurrences**
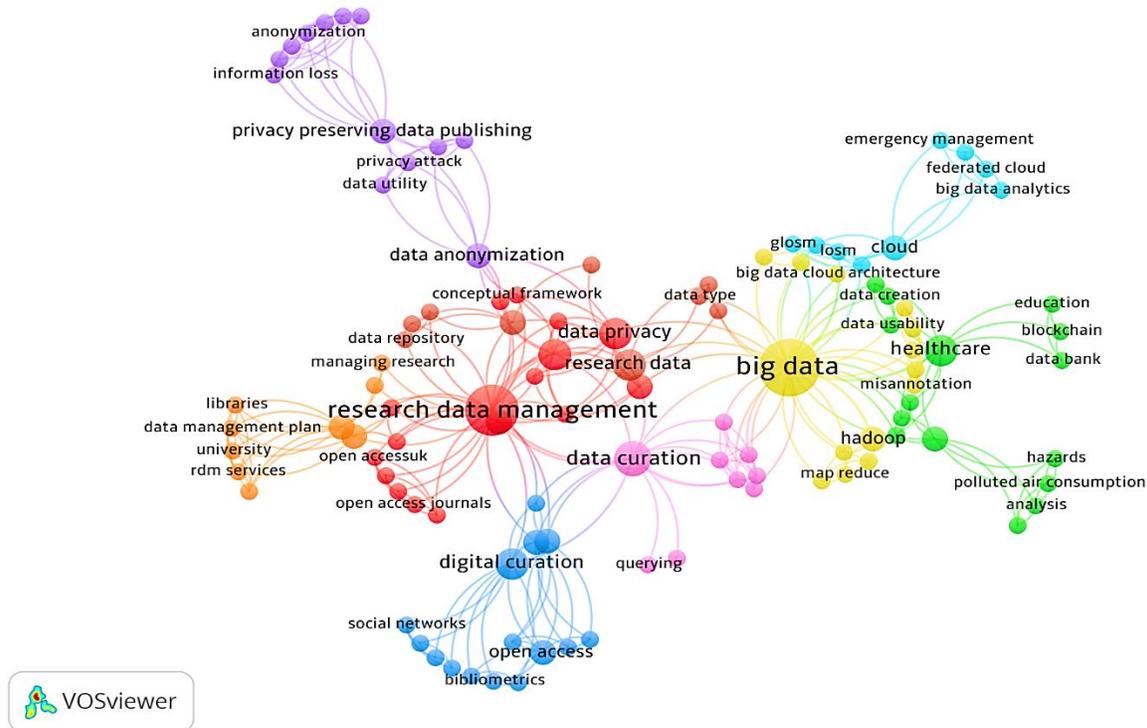
6.10. Indian Library and Information Science (LIS) Professionals' Contribution to RDM Research

RDM has become an emerging research interest for libraries. Libraries of developed countries have begun to provide a range of RDM services to assist researchers in supporting funding agencies and publishers' data requirements. These innovative practices and services are reflected in the scholarly articles, conference papers, research reports and working papers etc. In India, it is an emerging area wherein libraries have started exploring the impact of managing research data and data sharing through available policy guidelines. Figure 9 presents LIS vs other professionals' contributions to RDM research in India. Of a total of 136 publications, 38 (27.94%) were contributed by LIS professionals and the remaining 98 (72.05%) publications were contributed by professionals from other disciplines such as computer and information science, engineering, core and applied science, medical science, bioinformatics etc. LIS

professionals in India have explored major areas of RDM according to their needs and expertise, although they are yet to take the initiative to manage data generated by researchers. Studies (Anilkumar, 2018; Tripathi, and Sonker, 2017; Gajbe, et al., 2021) have been undertaken to assess the LIS professional's awareness about data curation, data archival policies, the infrastructure required, assessment of DMP tools technologies used and RDM services offered, found that Indian libraries' involvement in data management in research / academic institutes is still at a very early stage of development. The RDM services implemented by different universities' libraries in India for managing, organizing, curating and preserving research data generated at their campus indicates that libraries should take the lead in the curation of research data generated, besides curating and preserving, forming policies, collaborating with researchers, research support center, research supervisor, Lab managers, and members of the project team, IT center, administrators for research and development and other departments in order to achieve the research support data curation for research data access and use (Petrovich, 2020).

Figure 9. LIS *vs* Other Professionals Contribution to RDM Literature in India
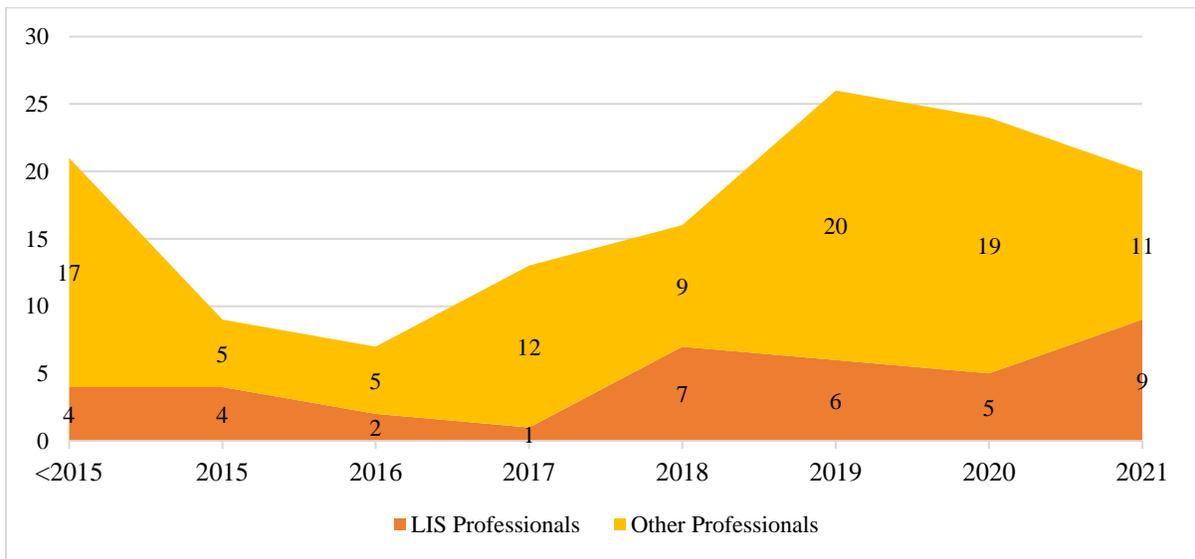
7   Conclusion

The current study was conducted to analyze published literature on RDM by Indian researchers indexed by the Scopus database. The results provide an overview of publication growth, citations, prolific authors, subject area of publications and top sources of publications. RDM published literature has been rapidly increasing since 2015. The study found that RDM related content also is published across several forms of scholarly literature like journal articles, conference papers, reviews, book chapters and editorial notes etc. The major research areas

within interdisciplinary subjects include "big data", *"research data management", "data curation", "data privacy", "Data Processing", "Research Data", "Data Repositories",* and *"Information Processing".* The top 5 highly cited works of literature have researched data curation, digital storage, data findings, data repository, and big data. Disciplines such as computer science, social sciences, biochemistry, genetics and molecular biology and mathematics are heavily engaged in the field of RDM. Study data indicates that the maximum collaboration of Indian researchers for RDM research is with developed countries like the United States and the United Kingdom. The study also states that Indian libraries' involvement in data management in research / academic institutes is still at a very early stage of development. This study can be taken as the beginning point for understanding the present RDM situation in India and can be explored further by researchers across other disciplines, institutions, and policymakers. Harnessing and sharing the power of research data will provide solutions to fight national and global challenges.

References

Anilkumar, N. (2018). Research data management in India: a pilot study. *The European Physical Journal Conferences, 186*. doi:10.1051/epjconf/201818603002

Azmi, N.I. & Abbas, S.M. (2018). Bibliometric Study of Publications on Drug Abuse in India From 2001 to 2016. *Library Progress (International)*, 38 *(1)*, 81-88, doi: 10.5958/2320-317X.2018.00008.9.

Bhakta, J., Kar, S. & Bhui, T. (2019). Bibliometric Mapping of LIS Research in India: A Study Seen Through the Mirror of Indian Citation Index. *Library Philosophy and Practice (e-journal),* 2349.

Bhardwaj, R. (2019). Research data management in higher educational institutions. *DESIDOC Journal of Library and Information Technology, 39*(6), 269-270. doi:10.14429/djlit.39.06.15281

Centre for Science and Technology Studies. (2020, 06 12). Retrieved from VOSviewer: https://www.vosviewer.com/

DST. (2012). *National Data Sharing and Accessibility Policy.* Retrieved from Department of Science and Technology, Government of India: https://dst.gov.in/sites/default/files/nsdi_gazette_0.pdf

Gajbe, Sagar Bhimrao, Amit, Tiwari, Gopalji & Singh, Ranjeet Kumar (2021), Evaluation and analysis of Data Management Plan tools: A parametric approach, *Information Processing & Management, 58*(3). doi: 10.1016/j.ipm.2020.102480.

Gupta, Nidhi, Arora, Surbhi & Chakravarty Rupak (2021). Science Mapping and Visualization of Research Data Management (RDM): Bibliometric and Scientometric Study, *Library Philosophy and Practice (e-journal),* 6096.

ICAR. (2014, October). *ICAR Guidelines for Internal Evaluation and Forwarding Research Papers to Scientific Journals and Data Management in ICAR Institutes.* Retrieved from

Indian Council of Agricultural Research: https://icar.org.in/files/ICAR-Guidelines-Research-Papers-2014.pdf

ICRISAT. (2014, March 05). *Data Management Policy.* Retrieved from International Crops Research Institute for the Semi-Arid Tropics: http://www.icrisat.org/wp-content/uploads/Data-management.pdf

Manu, T. (2018). Researchers' perceptions on research data management: a survey. *In International Conference on Exploring the Horizons of Library and Information Sciences: From Libraries to Knowledge Hubs.* Bangalore: Documentation Research and Training Centre (DRTC).

Oliveira, S. (2019). Bibliometric method for mapping the state-of-the-art and identifying research gaps and trends in literature: an essential instrument to support the development of scientific projects. In S. Kunosic, and E. mZerem, *Scientometrics Recent Advances.* London: IntechOpen.

Onyancha, O. (2016). Open research data in sub-Saharan Africa: a bibliometric study using the data citation index. *Publishing Research Quarterly*, 32, 227-246. doi: 10.1007/s12109-016-9463-6

Patel, D. (2016). Research data management: a conceptual framework. *Library Review, 65*, 226-241. doi:10.1108/LR-01-2016-0001

Pathak, M. & Bharat, K.A. (2020). Mapping ethnobotany research in India. *Ethnobotany Research and Applications*, 20, doi: 10.32859/era.20.49.1-12.

Payal, A. S., and 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. doi:10.14429/djlit.39.06.14451

Petrovich, G. (2020, 06 02). *A guide to research data management*. Retrieved from Lab Folder: https://www.labfolder.com/guide-research-data-management

Pritchard, A. (1969). Statistical bibliography or bibliometrics? *Journal of Documentation, 25*, 348–349.

Saeed, S., and Ali, P. (2019). Research data management and data sharing among research scholars of life sciences and social sciences. *DESIDOC Journal of Library and Information Technology, 39*(6), 290-299. doi:10.14429/djlit.39.06.14997

Selvi, M.G. & Gopalakrishnan, S.(2016). Indian Contribution on Crystallography: A Bibliometric Study Using Scopus Database. *Journal of Advances in Library and Information Science*, 5 *(4),* 390-399.

Singh, N., Monu, H., and Dhingra, N. (2018). Research data management policy and institutional framework. *Proceedings of IEEE 5th International Symposium on Emerging Trends and Technologies in Libraries and Information Services (ETTLIS 2018).* Noida: Bennett University.

Sjöstedt, E., Aldberg, H., and Jacobsson, C. (2015). *Guidelines for using bibliometrics at the Swedish Research Council.* Sweden: Swedish Research Council.

Surkis, A., and Read, K. (2015). Research data management. *J Med Libr Assoc., 103*(3), 154–156. doi:10.3163/1536-5050.103.3.011

Tripathi, D., and Pandy, S. (2018). Developing a conceptual framework of research data management for higher educational institutions. *Proceedings of IEEE 5th International*

Symposium on Emerging Trends and Technologies in Libraries and Information Services (ETTLIS 2018). Noida: Bennett University.

Tripathi, M., Kumar, S. & Babbar, P. (2018). Bibliometrics of social science and humanities research in India. *Current Science*, 114 *(11)*.

Tripathi, M. S., and Sonker, S. (2017). Research data management practices in university libraries: A study. *DESIDOC Journal of Library and Information Technology, 37*(6), 417-424. doi:10.14429/djlit.37.6.11336

Zhang, L., and Eichmann-Kalwara, N. (2019). Mapping the scholarly literature found in scopus on research data management: a bibliometric and data visualization approach. *Journal of Librarianship and Scholarly Communication, 7*.