

RESEARCH DATA MANAGEMENT IN LIBRARY: A BIBLIOMETRIC ANALYSIS OF THE PAST DECADE (2013-2023)

Zati Atiqah Mohamad Tanuri^{1*}
Abd Latif Abdul Rahman²
Zuraidah Arif³

¹ Universiti Teknologi MARA Cawangan Kedah, Kampus Sungai Petani; Email:zatiatiqah@uitm.edu.my;
ORCID ID (<https://orcid.org/0000-0003-0862-5999>)

² Universiti Teknologi MARA Cawangan Kedah, Kampus Sungai Petani; Email:ablatif@uitm.edu.my;
ORCID ID (<https://orcid.org/0000-0001-7970-2520>)

³ Universiti Teknologi MARA Cawangan Kedah, Kampus Sungai Petani; Email:zuraidah_arif@uitm.edu.my;
ORCID ID (<https://orcid.org/0000-0002-3357-1649>)

* Correspondence: Email:zatiatiqah@uitm.edu.my; 019-3698638.

Article history

Received date : 15-6-2023
Revised date : 16-6-2023
Accepted date : 28-7-2023
Published date : 13-9-2023

To cite this document:

Mohamad Tanuri, Z., Abdul Rahman, A. L., & Arif, Z. (2023). Research data management in library: A bibliometric analysis of the past decade (2013-2023). *Journal of Islamic, Social, Economics and Development (JISED)*, 8 (56), 19 - 31.

Abstract: *Research data management is an essential aspect of modern scientific research, and its importance has been increasingly recognized in recent years. With the growth of digital data and the demand for open science, the need for effective research data management practices has become more critical. Most academic research libraries are offering a variety of research data management services. The services are just expanded versions of conventional consulting or informative services, including assisting instructors and students in finding datasets or repositories. Conducting a bibliometric analysis of research data management can help address gap in the literature and provide a comprehensive overview of research trends in this field. The objectives of the paper are as follows: (1) to identify the research trends about research data management in library using bibliometric analysis; (2) to examines the clusters that include the co-occurrence of author keywords; (3) to produce a synthesised analysis of research trends on research data management in library, with particular attention to the geographic distribution of publications and authorship. This study reveals a bibliometric analysis of research data management in libraries as a theme in the current research trends and agenda. A total of 365 articles extracted from the Scopus database within the past 10 years, published from 2013 to 2023, were selected to be analyzed. The analysis of the academic work that was extracted throughout the search process was based on the following features; document and source type, subject area, most active source title, distribution of publications by countries, keyword analysis, and citation analysis.*

Keywords: *Research Data Management; Library; Bibliometric Analysis.*

Introduction

Research data management is an essential aspect of modern scientific research, and its importance has been increasingly recognized in recent years. With the growth of digital data and the demand for open science, the need for effective research data management practices has become more critical. Research data management (RDM) in libraries refers to the systematic and coordinated approach taken by libraries to support researchers in organizing, documenting, storing, preserving, sharing, and ensuring the long-term accessibility and usability of research data (Pryor, 2017; Cox, Verbaan, Sen, Devos, 2016; Corti, & Thompson, 2014). It encompasses the policies, infrastructure, services, and expertise provided by libraries to assist researchers in effectively managing their data throughout the research lifecycle.

Most academic research libraries are offering a variety of research data management services. The services are just expanded versions of conventional consulting or informative services, including assisting instructors and students in finding datasets or repositories. Libraries are becoming more involved with research data in a modest but rising number of ways, from aiding in data management plans to preparing and conserving research data for deposit in data repositories (Tenopir, Sandusky, Allard, Birch, 2014). Researchers and data professionals are primarily responsible for and in charge of RDM. Data librarians, data curators, and data stewards are part of the latter group (Koltay, 2023).

The literature in the context of libraries is so limited and recent, which may suggest that libraries are still unaware of their potential role in helping the research community (Hammad et al., 2021). Therefore, conducting a bibliometric analysis of research data management can help address this gap in the literature and provide a comprehensive overview of research trends in this field.

The objectives of the paper are as follows: (1) to identify the research trends about research data management in library using bibliometric analysis; (2) to examine the clusters that include the co-occurrence of author keywords; (3) to produce a synthesised analysis of research trends on research data management in library, with particular attention to the geographic distribution of publications and authorship. The remaining paper is divided into five sections that cover the methodology used in the current research, data analysis and results with the conclusion and suggestion.

Method & Material

The production, dissemination, and effect of scholarly publications are measured and evaluated using the quantitative research method known as bibliometric analysis. It entails gathering and analysing bibliographic information, such as citation counts, authorship, publication dates, and keywords, among other things. In academic research, bibliometric analysis is frequently employed, particularly in the domains of information science, bibliometrics, and scientometrics. Therefore, to identify and analyse the current growth of academic literature in the studies of research data management, bibliometric analysis was chosen as one of the analysis methods.

Scopus database was used to retrieve the relevant data. The Scopus database is used as the main source for data gathering in this study because Scopus is one of the largest curated abstract and citation databases with a wide global and regional coverage of scientific journals, conference proceedings, and books, while ensuring only the highest quality data are indexed through

rigorous content selection and re-evaluation by an independent Content Selection and Advisory Board (Baas, Schotten, Plume, Côté, Karimi, 2020).

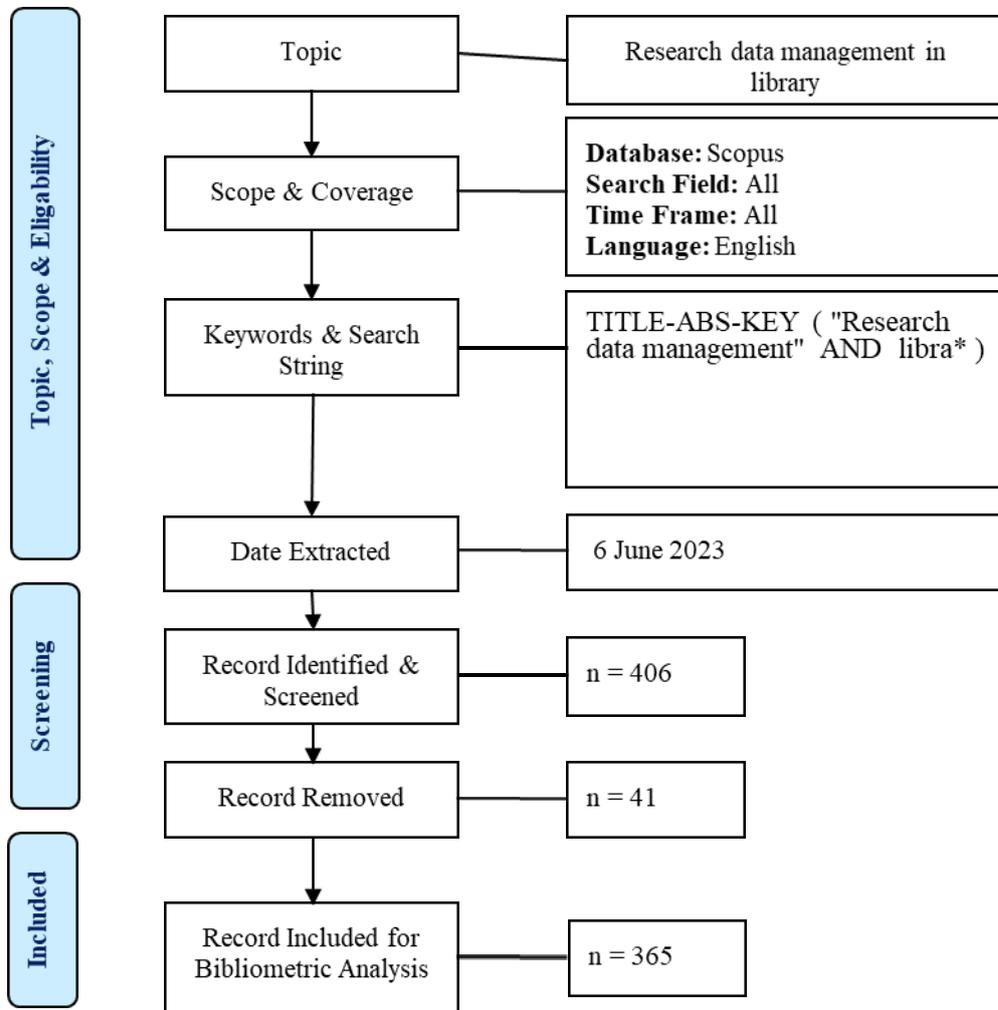


Figure 1: Flow Diagram of The Search Strategy

Source: Zakaria et al. (2020)

Data Extraction and Data Cleaning

In conducting the analysis, this study used search terms ("research data management" AND libra*) in article titles, abstracts, and keywords using wildcard symbols since using asterisk* will replace multiple characters (Ahmi, 2021). Any of these terms which appear in the title, abstract, or keywords of an article, are selected for bibliometric analysis. Using this method, a total of 406 documents were retrieved. The data extraction and cleaning process were summarised as follows: TITLE-ABS-KEY ("research data manage-ment" AND libra*) AND (EXCLUDE (PUBYEAR , 2012) OR EXCLUDE (PUBYEAR , 2011) OR EXCLUDE (PUBYEAR , 2010) OR EXCLUDE (PUBYEAR , 2009) OR EXCLUDE (PUBYEAR , 2008) OR EXCLUDE (PUBYEAR , 2003) OR EXCLUDE (PUBYEAR , 1979)) AND (LIMIT-TO (LANGUAGE , "English"))

After applying the above data extraction and cleaning processes, the final sample consisted of 365 articles that were considered sufficient for bibliometric analysis.

The bibliometric analysis for the current study was carried out using VOSviewer, a free tool for building and visualizing networks (vosviewer.com). This study also conducted using Harzing's publish or perish to retrieve and analyze scholarly citations. This study also looks into the influence of publications based on citation counts, impact per publication, and citation per publication using VOSviewer and publish or perish. Microsoft Excel was used to generate the editable tables and figure. All software used among the most well-known and frequently used for bibliometric data analysis.

Findings

The analysis of the academic work that was extracted throughout the search process was based on the following features; document and source type, subject area, most active source title, distribution of publication by countries, keywords analysis, and citation analysis.

Document and Source Types

Results of the document type, as presented in Table 1, show most of the studies on research adata management in library were published as an article (65.75%), followed by conference paper (16.99%), book chapter (6.85%), review (6.03%) and conference review (2.47%). Others were found as notes with (0.82%), book (0.55%), editorial and erratum (0.27%).

Table 1: Document Type

Document Type	Total Publications (TP)	Percentage (%)
Article	240	65.75%
Conference Paper	62	16.99%
Book Chapter	25	6.85%
Review	22	6.03%
Conference Review	9	2.47%
Note	3	0.82%
Book	2	0.55%
Editorial	1	0.27%
Erratum	1	0.27%
Total	365	100.00

The study found that these materials are divided into four different source types: books, journals, book series, and conference proceedings (see Table 2). Out of the 365 documents, 73.42% have been published in journals, 11.78% have been published in conference proceedings, and 7.95% have been published in book series. On the other hand, with 6.85%, book had the lowest publication.

Table 2: Source Type

Source Type	Total Publications (TP)	Percentage (%)
Journal	268	73.42%
Conference Proceeding	43	11.78%
Book Series	29	7.95%
Book	25	6.85%
Total	365	100.00

Year of Publications/Evolution of Published Studies

The year 2019 was the highest year for publication in this area, with the total number of the articles published 52 (14.25%). Followed by 2021 (13.42%) and 2020 (13.15%). Meanwhile, in the year 2018, less than 10% of publications were produced regarding this topic or theme (Table 3).

Table 3: Year of Publications

Year	Total Publications	Percentage (%)
2023	18	4.93%
2022	37	10.14%
2021	49	13.42%
2020	48	13.15%
2019	52	14.25%
2018	41	11.23%
2017	43	11.78%
2016	27	7.40%
2015	19	5.21%
2014	22	6.03%
2013	9	2.47%
Total	365	100.00

Documents by year



Fig. 2: Document by Year

Based on the pattern and growth, it seems the number of documents published yearly is gradually declining.

Subject Area

Table 4 present the distribution of publications across various subject areas. The analysis revealed that highest number of publication were under field of social sciences with a total 273 (74.79%) publications. This followed by “computer science (34.25%). The distribution of the publications across other subject areas was relatively smaller which is below 10% of publication, including arts and humanities, mathematics, medicine, engineering, business management and accounting, health professions, decision sciences, biochemistry, genetics and molecular biology, agricultural and biological sciences, multidisciplinary, earth and planetary

sciences, economics, econometrics and finance, chemistry environmental science, chemical engineering, materials science, neuro science, pharmacology, toxicology and pharmaceuticals, physics and astronomy.

Table 4: Subject Area

Subject Area	Total Publications	Percentage (%)
Social Sciences	273	74.79%
Computer Science	125	34.25%
Arts and Humanities	30	8.22%
Mathematics	26	7.12%
Medicine	23	6.30%
Engineering	19	5.21%
Business, Management and Accounting	11	3.01%
Health Professions	9	2.47%
Decision Sciences	7	1.92%
Biochemistry, Genetics and Molecular Biology	6	1.64%
Agricultural and Biological Sciences	5	1.37%
Multidisciplinary	5	1.37%
Earth and Planetary Sciences	4	1.10%
Economics, Econometrics and Finance	4	1.10%
Chemistry	3	0.82%
Environmental Science	3	0.82%
Chemical Engineering	1	0.27%
Materials Science	1	0.27%
Neuroscience	1	0.27%
Pharmacology, Toxicology and Pharmaceutics	1	0.27%
Physics and Astronomy	1	0.27%

Most Active Source Title

Table 5 presents the most active source titles on library research data management. The highest source was Lecture Notes in Computer Science Including Subseries Lecture Notes In with 18 (4.93%) publications. The second and third highest were Journal of Academic Librarianship and IFLA Journal with 16 and 12 publications. Followed by Data Science Journal, Ceur Workshop Proceedings with the Library Management and New Review of Academic Librarianship with total 9 (2.47%). Whereas ACM International Conference Proceeding Series and Advances in Library Administration and Organization was the lowest source with 3 (0.82%) publications.

Table 5: Most Active Source Title

Source Title	Total Publications	Percentage (%)
Lecture Notes In Computer Science Including Subseries Lecture Notes In Artificial Intelligence And Lecture Notes In Bioinformatics	18	4.93%
Journal Of Academic Librarianship	16	4.38%
IFLA Journal	12	3.29%
Data Science Journal	11	3.01%
Ceur Workshop Proceedings	9	2.47%
Liber Quarterly	9	2.47%
Library Management	9	2.47%

Source Title	Total Publications	Percentage (%)
New Review Of Academic Librarianship	9	2.47%
Journal Of The Medical Library Association	8	2.19%
Library Philosophy And Practice	8	2.19%
Proceedings Of The Association For Information Science And Technology	8	2.19%
Communications In Computer And Information Science	7	1.92%
Desidoc Journal Of Library And Information Technology	6	1.64%
Health Information And Libraries Journal	6	1.64%
Issues In Science And Technology Librarianship	6	1.64%
Journal Of Librarianship And Scholarly Communication	6	1.64%
Evidence Based Library And Information Practice	5	1.37%
GI Conference Series Conference Proceedings	5	1.37%
Global Knowledge Memory And Communication	5	1.37%
Information Services And Use	5	1.37%
Journal Of Librarianship And Information Science	5	1.37%
Library Hi Tech News	5	1.37%
Libri	5	1.37%
Proceedings Of The ACM IEEE Joint Conference On Digital Libraries	5	1.37%
Cases On Research Support Services In Academic Libraries	4	1.10%
College And Research Libraries	4	1.10%
Electronic Library	4	1.10%
Library And Information Science Research	4	1.10%
Medical Reference Services Quarterly	4	1.10%
Program	4	1.10%
ACM International Conference Proceeding Series	3	0.82%
Advances In Library Administration And Organization	3	0.82%

Keyword analysis

Keyword from the documents gathered were analysed using the full counting method via VOSviewer software, each of which had a minimum of 5 occurrences. The colour, circle size, font size, and thickness of the connecting lines represent the relationships with other keywords. For example, keywords with the same colour are commonly listed. Thus, in this study, research data management, metadata, research data services, data sharing, data repositories and data management had similar colours, suggesting that these keywords are closely related and usually occur together.

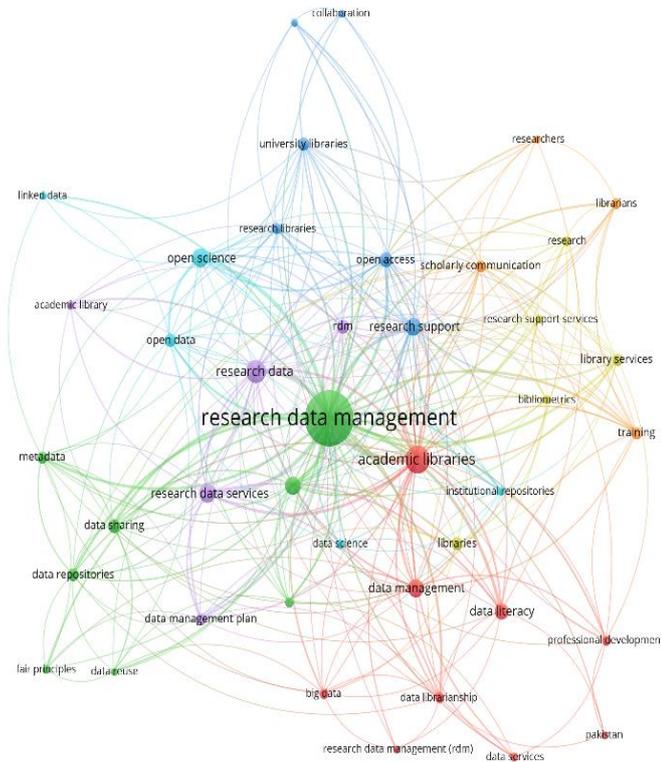


Figure 3: Network visualization map of the author keywords

The table presents a bibliometric analysis of keywords related to library research data management, based on the total number of publications in which these keywords appeared. The latest citations indicates that research data management is the most frequently studied subject area, appearing in 652.88% of the publications. Information management is the second most popular keyword, with a frequency of 27.40%. Other popular keywords include academic libraries (16.44%), research data (15.34%), and libraries (11.51%). The data shows that a variety of research topics have been investigated in relation to research data management, such as open science, research data services and data curation.

Table 6: Top Keywords

Author Keywords	Total Publications	Percentage (%)
Research Data Management	193	52.88%
Information Management	100	27.40%
Research Data Managements	89	24.38%
Academic Libraries	60	16.44%
Research Data	56	15.34%
Libraries	42	11.51%
Digital Libraries	38	10.41%
Human	32	8.77%
Data Management	30	8.22%
Library	27	7.40%
Article	24	6.58%
Open Science	24	6.58%

Research Data Services	24	6.58%
Data Curation	23	6.30%
Information Services	22	6.03%
Research Support	22	6.03%
Librarian	20	5.48%
Data Repositories	19	5.21%
Data Sharing	19	5.21%
Humans	19	5.21%

Geographical Distribution of Publications - Most Influential Countries

This study also presents the most active countries that have published documents on library research data management. Table 7 records that the top countries that contributed to the publications were the United States (32.33%) with 118 publications, followed by United Kingdom (10.14%), and the third top country is India (5.75%).

Table 7: Top 20 Countries contributed to the publications

Country	Total Publications	Percentage (%)
United States	118	32.33%
United Kingdom	37	10.14%
India	21	5.75%
Germany	20	5.48%
Australia	19	5.21%
South Africa	14	3.84%
Canada	13	3.56%
China	12	3.29%
Pakistan	11	3.01%
Portugal	11	3.01%
Nigeria	9	2.47%
Hungary	7	1.92%
Netherlands	7	1.92%
Italy	6	1.64%
Austria	5	1.37%
Spain	5	1.37%
Sweden	5	1.37%

Citation Analysis

To identify the citation metric for the data extracted from the Scopus database, Harzing's Publish or Perish software was applied. The citation metrics for the publications that were retrieved as of June 6, 2023 are shown in Table 8. The number of citations together with their citations per year, citations per work, and citations per author are included in the brief description.

There were 365 papers total, with 3328 citations, average 332.80 citations year in the field of Research data management in library. Each papers received 9.12 citations, and the overall h-index and g index for all publications were 27 and 45, respectively.

Table 8: Citations Metrics

Metrics	Data
Publication years	2013-2023
Citation years	10 (2013-2023)
Papers	365
Citations	3328
Citations/year	332.80
Citations/paper	9.12
Papers/author	2.55
h-index	27
g-index	45

The top 20 most cited articles on research data management in library are presented in Table 9 below. C. Tenopir, R.J. Sandusky, S. Allard, B. Birch, with the title “Research data management services in academic research libraries and perceptions of librarians”, was the most cited article to date, with 157 total citations altogether. Followed by S. Corral, M.A. Kennan, W. Afzal with the article entitled “Bibliometrics and research data management services: Emerging trends in library support for research”.

Table 9: Highly cited articles

No.	Authors	Title	Year	Cites	Cites per Year
1	C. Tenopir, R.J. Sandusky, S. Allard, B. Birch	Research data management services in academic research libraries and perceptions of librarians	2014	157	17.44
2	S. Corral, M.A. Kennan, W. Afzal	Bibliometrics and research data management services: Emerging trends in library support for research	2013	155	15.5
3	V. Sharma, J. Eckels, G.K. Taylor, N.J. Shulman, A.B. Stergachis, S.A. Joyner, P. Yan, J.R. Whiteaker, G.N. Halusa, B. Schilling, B.W. Gibson, C.M. Colangelo, A.G. Paulovich, S.A. Carr, J.D. Jaffe, M.J. Maccoss, B. Maclean	Panorama: A targeted proteomics knowledge base	2014	152	16.89
4	A.M. Cox, S. Pinfield	Research data management and libraries: Current activities and future priorities	2014	140	15.56
5	A.M. Cox, M.A. Kennan, L. Lyon, S. Pinfield	Developments in research data management in academic libraries: Towards an understanding of research data service maturity	2017	110	18.33

6	S. Pinfield, A.M. Cox, J. Smith	Research data management and libraries: Relationships, activities, drivers and influences	2014	97	10.78
7	T. Koltay	Data literacy for researchers and data librarians	2017	69	11.5
8	A. Yoon, T. Schultz	Research data management services in academic libraries in the US: A content analysis of libraries' websites	2017	61	10.17
9	R.C. Amorim, J.A. Castro, J. Rocha da Silva, C. Ribeiro	A comparison of research data management platforms: architecture, flexible metadata and interoperability	2017	57	9.5
10	T. Koltay	Data governance, data literacy and the management of data quality	2016	54	7.71
11	A.M. Cox, M.A. Kennan, L. Lyon, S. Pinfield, L. Scaffi	Maturing research data services and the transformation of academic libraries	2019	52	13
12	R.A. Brown, M. Wolski, J. Richardson	Developing new skills for research support librarians	2015	47	5.88
13	A.L. Whitmire, M. Boock, S.C. Sutton	Variability in academic research data management practices: Implications for data services development from a faculty survey	2015	44	5.5
14	A. Surkis, K. Read	Research data management	2015	40	5
15	M.A. Kennan, S. Corral, W. Afzal	“making space” in practice and education: Research support services in academic libraries	2014	40	4.44
16	J. Cox	Positioning the Academic Library within the Institution: A Literature Review	2018	38	7.6
17	L. Perrier, E. Blondal, A.P. Ayala, D. Dearborn, T. Kenny, D. Lightfoot, R. Reka, M. Thuna, L. Trimble, H. MacDonald	Research data management in academic institutions: A scoping review	2017	37	6.17
18	E. Verbaan, A.M. Cox	Occupational Sub-Cultures, Jurisdictional Struggle and Third Space: Theorising Professional Service Responses to Research Data Management	2014	37	4.11

19	A.M. Cox, W.W.T. Tam	A critical analysis of lifecycle models of the research process and research data management	2018	36	7.2
20	L. Federer	Research data management in the age of big data: Roles and opportunities for librarians	2016	36	5.14

Conclusion

Modern scientific research requires the effective management of research data, and this need has grown in prominence in recent years. Effective research data management procedures are more important than ever due to the expansion of digital data and the demand for open science. Research trends and gaps can be found and the current condition of the subject can be determined with the use of a bibliometric examination of research data management.

This study reveals a bibliometric analysis of research data management in libraries as a theme in the current research trends and agenda. A total of 365 articles extracted from the Scopus database within the past 10 years, published from 2013 to 2023, were selected to be analyzed. The analysis of the academic work that was extracted throughout the search process was based on the following features; document and source type, subject area, most active source title, distribution of publications by countries, keyword analysis, and citation analysis. Results indicate that keywords such as "research data management, metadata, research data services, data sharing, data repositories," and data management are closely related and usually occur together. Also, the most significant subject areas for publications are social science and computer science. The result also indicates declining numbers of publications starting in 2022.

The limitation of the study is that some search criteria were created solely using publications found in the Scopus database. As a result, some pertinent articles published in journals other than Scopus or that did not satisfy the search parameters may have been missed. For future research, the study can be extended by using other databases, such as Web of Science (WoS) or Google Scholar.

References

- Ahmi, A., & Mohamad, R. (2019). Bibliometric Analysis of Global Scientific Literature on Web Accessibility. *International Journal of Recent Technology and Engineering*, 7(6), 250–258.
- Ahmi, A., & Mohd Nasir, M. H. (2019). Examining the Trend of the Research on eXtensible Business Reporting Language (XBRL): A Bibliometric Review. *International Journal of Innovation, Creativity and Change*, 5(2), 1145–1167.
- Baas, J. Schotten, M., Plume, A., Côté, G. Karimi, R. (2020)., Scopus as a curated, high-quality bibliometric data source for academic research in quantitative science studies. *Quantitative Science Studies* 2020; 1 (1): 377–386. doi: https://doi.org/10.1162/qss_a_00019
- Corti, L., & Thompson, P. (2014). *Managing and sharing research data: A guide to good practice*. Sage Publications.
- Cox, A. M., Verbaan, E., Sen, B., & Devos, R. (2016). How researchers manage their data: Insights from the academic community. *PLOS ONE*, 11(4), e0152971. doi:10.1371/journal.pone.0152971

- Hamad, F., Al-Fadel, M. and Al-Soub, A. (2021). Awareness of research data management services at academic libraries in Jordan: roles, responsibilities and challenges. *New Review of Academic Librarianship*, 27 (1), 76-96, doi: 10.1080/13614533.2019.1691027
- Pryor, G. (2017). Research data management: Libraries' role in supporting faculty data sharing behaviors. *The Journal of Academic Librarianship*, 43(5), 454-460. doi:10.1016/j.acalib.2017.07.006
- Tenopir, C., & Sandusky, R. (2017). Research data management services in academic libraries in the US: A content analysis of libraries' websites. *Journal of the Association for Information Science and Technology*, 68(2), 421-433. doi:10.1002/asi.23644
- Koltay, T. (2023). *Encyclopedia of Data Science and Machine Learning*. Doi:10.4018/978-1-7998-9220-5.ch072
- Zakaria, R., Ahmi, A., Ahmad, A. H., & Othman, Z. (2020) Worldwide melatonin research: A bibliometric analysis of the published literature between 2015 and 2019, *Chronobiology International*, 38(1), 27-37. <https://doi.org/10.1080/07420528.2020.1838534>