Record and Library Journal
https://e-journal.unair.ac.id/index.php/RLJ

Volume 11, No. 1, 2025
e-ISSN: 2442-5168

Knowledge mapping of research data management: Uncovering themes and developments through co-occurrence and bibliometric analysis

Mohd Ikhwan Ismail[1,2], A. Abrizah[2] and Samsul Farid Samsuddin[2]

[1]Hamzah Sendut Library, Universiti Sains Malaysia, 11800 Penang, Malaysia
[2]Department of Library & Information Science, Faculty of Arts and Social Sciences, Universiti Malaya, 50603 Kuala Lumpur, Malaysia

Paper Type:

Literature Review

Abstract

Background of the study: Discussions on RDM have grown rapidly in scholarly platforms, emerging as a key topic within library and information science (LIS). While existing studies have reviewed and analyzed RDM literature, their scope is often limited to specific areas or timeframes. It is necessary for a detailed and current analysis of RDM literature, providing deeper insights into its complexities, evolution, and future directions.

Purpose: The study presents mapping knowledge domains as a method to uncover the thematic landscape, identify significant clusters, and provide a structured understanding of interconnected concepts within the field of RDM.

Method: Data were retrieved from Elsevier's Scopus database as of August 2023. The study conducts bibliometric analysis to examine geographical distribution, publication outlet, authorship trends, and performance metrics within the field.

Findings: The dataset spans from 1977 to 2023, with an increase in publications exceeding ten per year from 2012 onwards, amounting to 684 documents in various languages and reference types. The study identifies four research clusters derived from these documents, highlighting key themes namely, RDM services, data sharing, information systems, and data management.

Conclusion: The findings underscore the growth of RDM-related research and contribute to a deeper understanding of the underlying structure of RDM, for researchers, practitioners, and policymakers, enabling them to address current challenges and anticipate future developments in this rapidly evolving field.

Keywords: Research Data Management (RDM), Bibliometric analysis, Theme identification, Literature growth, Co-occurrence analysis

Submitted: 27 October 2024
Revised: 17 February 2025
Accepted: 12 May 2025
Online: 25 June 2025

* Correspondence:
Mohd Ikhwan Ismail

E-mail:
ikhwanismail@usm.my

Ismail, M. I., Abrizah, A., & Samsuddin, S. F. (2025). Knowledge mapping of research data management: Uncovering themes and developments through co-occurrence and bibliometric analysis. *Record and Library Journal, 11*(1), 234-258. DOI: 10.20473/rlj.V11-I1.2025.234-258.

Introduction

Research data management (RDM), an integral part of the research process, has gained widespread attention in empirical studies over the past decade, largely driven by the growing emphasis on open science. Calls for effective RDM practices emphasize the need to align seamlessly with principles of openness and transparency in research. RDM, which concerns the organization of data from entry into the research lifecycle through storage, dissemination, and archiving of valuable results, is crucial for various stakeholders, including government and funders, university leadership, research management/support units, and researchers (Flores et al., 2015). The creation of RDM services in research organizations is influenced by various drivers and factors. Perspectives of RDM could be derived from different components related to strategies, policies, guidelines, processes, technologies and services (Pinfield et al., 2014). The development of RDM would be interrelated with the stakeholders, drivers, and influencing factors.

As institutions and funding bodies increasingly acknowledge the significance of effective RDM, there has been a substantial rise in the development of policies regulating data management practices across research performing organizations (RPOs) globally. In developed countries such as the United States, the United Kingdom, and Australia, RDM policies continue to evolve, with universities implementing varying frameworks, ranging from partial to comprehensive approaches to strengthen data governance (Liu et al., 2020). This transition aligns with a broader trend of digitalization in academia, where researchers are expected to take on greater responsibility for data management in compliance with mandates established by funding agencies and major academic publishers (Wilms et al., 2018). Moreover, recent studies suggest that the effective implementation of RDM not only strengthens research outcomes by enhancing data accessibility and reproducibility but also enhances the visibility and societal impact of research outputs (Andrikopoulou et al., 2022; Cruz et al., 2019; Mosha & Ngulube, 2023; Xu, 2022). While developed countries have shown the way, developing countries are also recognizing the importance of RDM. A recent study in Malaysia has emphasized the importance of focusing on tasks related to data governance roles, areas, and decision domains within RPOs, highlighting their critical connection to the effective implementation of RDM practices Hazmi et al. (2023). The successful adoption of RDM practices also requires robust data stewardship capacity-building initiatives in RPOs to equip researchers, institutions, and stakeholders with the necessary skills, infrastructure, and resources Ismail et al. (2024). As a result, the establishment of comprehensive RDM strategies is essential for fostering a research environment in which data stewardship is prioritized, ultimately contributing to the advancement of scientific inquiry and supporting broader institutional and national development goals (Zotoo & Liu, 2019).

Given the rapid development in RDM practices and their increasing importance in RPOs, analyzing the growth and evolution of RDM literature has become essential. Bibliometric analysis has emerged as a critical tool in academic research for systematically evaluating the evolution and trends within a specific field, thereby providing insights into the scholarly landscape. This method quantitatively analyzes publications by measuring various metrics such as citation counts, keyword frequencies, author collaborations, and publication sources (Fadhli et al., 2023; Us & Gerulaitienė, 2022). The application of bibliometrics, augmented by knowledge visualization techniques, highlights the methodology's effectiveness in providing an objective overview of research developments. By analyzing bibliographic data, it enables the extraction of meaningful insights into thematic focus and trends (Masalimova et al., 2023). In addition, the versatility of bibliometric techniques (Verma & Gustafsson, 2020), underscores their applicability across a wide range of academic disciplines, thereby affirming their significance in fostering a comprehensive understanding of research trajectories. Overall,

Ismail, M. I., Abrizah, A., & Samsuddin, S. F. (2025). Knowledge mapping of research data management: Uncovering themes and developments through co-occurrence and bibliometric analysis. *Record and Library Journal, 11*(1), 234-258. DOI: 10.20473/rlj.V11-I1.2025.234-258.

bibliometric analysis serves as a navigational tool for exploring complex scientific outputs and an evaluative framework that promotes scholarly collaboration and drives research advancements in emerging fields. Therefore, this study leverages bibliometric tools to map the landscape of RDM research, offering insights into its growth, impact, and emerging directions.

Literatur Review

Discussions on RDM have rapidly expanded across scholarly platforms, making it a popular research topic in library and information science (LIS) literature. While several studies have systematically reviewed and conducted bibliometric analyses on RDM literature, they are often limited to specific areas. For example, Perrier et al. (2017), who conducted a scoping review on RDM in academic institutions covering 301 articles published from 1995 to 2016, categorized RDM into five distinct focus areas: stakeholder, data, library, tool/device, publication, and data quality. Ashiq et al. (2022), who explored RDM literature published from 2016 to 2020, specifically delved into the challenges, services, skills, and factors associated with RDM practices among researchers, and the services provided by academic libraries. Zhang & Eichmann-Kalwara (2019) utilized CiteSpace for data visualization in their bibliometric analysis of RDM literature within the Scopus database. Examining documents published between 1945 and 2018, the study identified six prominent RDM themes: scientific collaboration, research support service, data literacy, knowledge manager, organizational environment, and information literacy. Additionally, the research highlighted significant research clusters, specifically in scientific collaboration, research support service, and data literacy. The study also observed a noteworthy increase in LIS topics, encompassing digital libraries, big data, data sharing, and data curation. Gupta et al. (2021) focused on publications within the Web of Science (WoS) Core Collection database spanning from 1989 to 2021, with the aim of elucidating the current state of RDM literature. Utilizing keyword co-occurrence mapping, the study unveiled six distinct clusters in WoS from RDM publications: data management plan (DMP); data management; data literacy and lifecycle approach; privacy and quality of data; data science and literacy; and data preservation options and sharing. Nevertheless, there exists a gap in the comprehensive profiling of RDM literature in these studies, impeding a thorough understanding of the knowledge structure and hindering the assessment of emerging trends and future research focuses.

Several bibliometric studies have explored the research domains associated with RDM publications. Naseema & Sevukan (2022) employed three main tools for bibliometric analysis, including the R package Biblioshiny, ScientoPy, and VOSviewer, to analyze RDM publications indexed in Scopus and the Web of Science. The study's key findings indicate that the most prominent research domains in RDM are Computer Science and Library and Information Science. This is similar to the study by Preeti & Amit Kumar (2023), which found that the majority of RDM publications are categorized under Computer Science, although their analysis was limited to publications indexed in Scopus.

In a recent study, Shah et al. (2024) examined RDM publications indexed in the Web of Science from 2001 to 2021 without restrictions on document type. The study utilized various visualization tools such as VOSviewer, Biblioshiny, ScientoPy and MS Excel. Notably, the findings revealed that librarians and information scientists were the primary contributors to RDM research, with Sheffield and Pittsburgh identified as the most productive institutions in this field. Furthermore, the study highlighted that Information Science and Library Science were the dominant research domains for RDM publications. Another recent study by Lee et al. (2024) analyzed 403 RDM publications indexed in the Web of Science. One of the key findings of the study indicated that RDM research domains are shaped by established standards and exhibited varying trends across countries, reflecting their specific interests and circumstances.

Ismail, M. I., Abrizah, A., & Samsuddin, S. F. (2025). Knowledge mapping of research data management: Uncovering themes and developments through co-occurrence and bibliometric analysis. *Record and Library Journal, 11*(1), 234-258. DOI: 10.20473/rlj.V11-I1.2025.234-258.

Objectives

Additional study on previous research is necessary for a detailed and current analysis of RDM literature, providing deeper insights into its complexities, evolution, and potential future directions. This study centers on the analysis of keyword co-occurrences within a network and utilizes the mapping of knowledge domains. It aims to develop clusters from RDM literature, with the purpose of identifying subtopics or thematic areas within the field. The keywords co-occurrence network is a valuable tool for identifying research areas (Liu & Mei, 2016; Zhu & Zhang, 2020). In this study, mapping knowledge domains aims to characterize RDM as an emerging interdisciplinary field, focusing on the process of mining, analyzing, sorting, enabling navigation of, and displaying knowledge (Shiffrin et al., 2004). This concept of mapping knowledge domains could also be called knowledge graphs or visualization within the context of social network analysis (Zhu et al., 2015). This analysis could reveal that the hidden connections between the publications and the mapping knowledge domains cannot be easily interpreted if done manually (Li et al., 2017). Therefore, this study presents mapping knowledge domains as a power tool to reveal the thematic landscape, identify significant clusters, and offer a structured understanding of interconnected concepts within the field of RDM (Stopar & Bartol, 2019). The study addresses the following research question: "*How can the relationships among research data management literature be described and analyzed in a representative, dynamic, and scalable way?"*

Method

Data for this study were retrieved from Elsevier's Scopus database as of August 2023. Scopus has been chosen because it is one of the largest citation databases of peer-reviewed literature, with strength in inter-disciplinary features and the nature of RDM as a research field (Zhang & Eichmann-Kalwara, 2019). Scopus covered various types of research publications, ensuring the quality of data indexed through meticulous data selection and re-evaluation by the Scopus Content Selection and Advisory Board (Baas et al., 2020). The use of the Scopus database is justified due to its broad coverage, global accessibility, longitudinal data, interdisciplinary content and data quality. This inclusivity allows the study to access a diverse array of publications, providing a comprehensive overview of research outputs (Baas et al., 2020; Martín-Martín et al., 2021). These features collectively establish a robust foundation for the study, facilitating a thorough and nuanced analysis of the research landscape (Wahid et al., 2020).

The study employs a set of inclusive criteria, encompassing publications in various languages and source types, and document types to ensure a diverse and comprehensive representation of the RDM literature. Initial steps involve the identification and selection of keywords relevant to RDM for subsequent analysis. The search in the Scopus database utilizes the keyword "research data management" in the title, abstract, and keywords fields, with an additional exact keyword "research data management" in the keyword search. The search was executed on August 17, 2023, as illustrated in Figure 1, depicting the literature search flow. A total of 684 records were identified, and all are deemed usable. They have been deposited in OpenData@USM, accessible at https://opendata.usm.my/handle/123456789/74720.

Following the literature search, publication data were directly extracted from the Scopus database using two export data formats: (i) RIS and (ii) CSV. The identified publications from these export types undergo analysis using three application tools: (a) VOSviewer version 1.6.19 for generating and visualizing bibliometric networks, including keywords co-occurrence networks for the RIS data type; (b) Microsoft Excel for retrieving frequencies and percentages of publications and developing graphs and tables for the CSV data type; and (c) Harzing's Publish or Perish on Microsoft Windows version 8.9.4538 for generating citation information.

Ismail, M. I., Abrizah, A., & Samsuddin, S. F. (2025). Knowledge mapping of research data management: Uncovering themes and developments through co-occurrence and bibliometric analysis. *Record and Library Journal, 11*(1), 234-258. DOI: 10.20473/rlj.V11-I1.2025.234-258.

Subsequently, the clusters within the 684 RDM publications were identified through co-occurrence analysis, representing cohesive groups of keywords and topics within RDM literature. Each cluster is characterized based on prevalent themes and keywords, offering insights into the diverse aspects of RDM. The study utilizes visualizations, such as network maps and graphs, to present co-occurrence patterns, geographical distribution, and cluster relationships. Quantitative metrics assess the growth of RDM-related research over time, including an analysis of annual publication counts with a focus on identifying significant trends and inflection points. The study conducts a geographical analysis to explore the global distribution of RDM research, identifying countries or regions with notable contributions to the field. Authorship trends are examined to understand the key contributors to RDM literature, with a specific focus on identifying prolific authors and collaborative networks. Additionally, performance metrics, such as citation counts, are considered to evaluate the impact and visibility of RDM publications. This comprehensive methodology aims to provide a detailed and nuanced exploration of the knowledge landscape in the realm of RDM (van Eck et al., 2017).

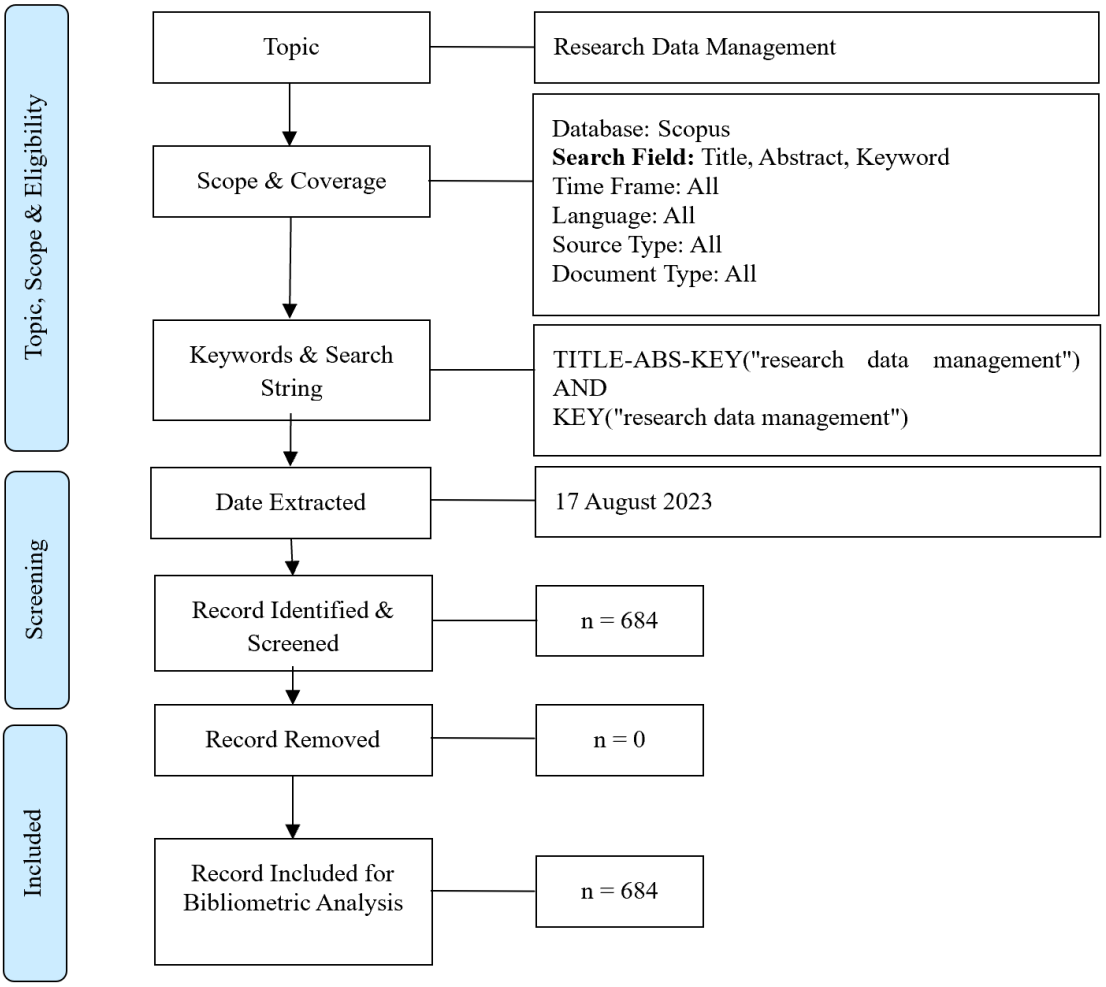


Figure 1. Flowchart illustrating the literature search and data analysis process in the study

Ismail, M. I., Abrizah, A., & Samsuddin, S. F. (2025). Knowledge mapping of research data management: Uncovering themes and developments through co-occurrence and bibliometric analysis. *Record and Library Journal, 11*(1), 234-258. DOI: 10.20473/rlj.V11-I1.2025.234-258.

Result and Discussion

RDM Topics

A total of 160 keywords, encompassing both author and index keywords, were identified from the 684 Research Data Management (RDM) publications. These keywords collectively appeared 3,937 times across the analyzed 684 publications. The study identified issues with similar keywords differing in singular and plural terms, such as "Research Data Management" and "Research Data Managements," where the singular term is preferred for proper nouns, and the plural term is used for common nouns. Certain keywords exhibited synonymous meanings or connotations but were expressed through different terms or spellings, for instance, "Fair" and "Fair Principles." Consequently, the study addressed these variations by combining certain keywords, creating unique terms through the thesaurus method in VOSviewer to ensure accuracy in the results. Following this cleaning process, the study identified and utilized 145 keywords as descriptors, aiding in the identification of the subject matter within the publications (Lozano et al., 2019).

Table 1 shows the top 20 keywords from the utilized 145 keywords in the 684 publications. Each publication may be associated with one or more keywords. The most frequently occurring keywords include broad terms such as "Research Data Management," "Information Management," and "Research Data." "Libraries" appears 4th in the list of top keywords used. Other keywords used are associated with concepts and areas in the management and delivery of digital content, for example, "Metadata", "Data Repository", "Data Curation", "Digital Storage" and "Information Services". RDM is closely linked to openness and transparency, substantiated by statements made in the paper's introduction, as reflected in keywords like "Open Science", "Data Sharing", "Open Data", "Open Access" and "FAIR Principles".

Table 1. Top 20 Keywords from 160 Keywords in 684 Publications

No.	Keywords	Total number appearing in publications
1	Research Data Management	877
2	Information Management	346
3	Research Data	138
4	Libraries	89
5	Metadata	84
6	Open Science	76
7	Data Sharing	66
8	Open Data	52
9	Data Management	48
10	Data Curation	47
11	Data Repositories	43
12	Digital Storage	43
13	Digital Libraries	37
14	Information Processing	36
15	Information Services	34
16	FAIR Principles	33
17	Life Cycle	33
18	Semantics	32
19	Open Access	31
20	Big Data	30

* The total number of appearing in publications for each keyword could be more than the actual number after merging some keywords

Ismail, M. I., Abrizah, A., & Samsuddin, S. F. (2025). Knowledge mapping of research data management: Uncovering themes and developments through co-occurrence and bibliometric analysis. *Record and Library Journal, 11*(1), 234-258. DOI: 10.20473/rlj.V11-I1.2025.234-258.

RDM Research Areas

Research areas or clusters were identified based on the 145 keywords extracted from RDM publications in this study. The clusters were developed with a minimum threshold of 15 occurrences for keywords, excluding those keywords with low occurrences. Based on this threshold, 53 keywords were selected through this process and visualized based on their respective clusters using VOSviewer.

Figure 2 shows the network visualization of the keywords' co-occurrence network. In this representation, the circle size, font size, colour, and thickness of the connecting lines indicate the link strength of the relationship between keywords. The network indicates that keywords of the same colour are commonly listed together, indicating their close association. Four clusters emerged from the keywords co-occurrences network, categorized as follows in themes: "RDM services" (Cluster 1, red, 18 keywords); "Data sharing" (Cluster 2, dark green, 8 keywords); "Information systems" (Cluster 3, blue, 12 keywords) and "Data management" (Cluster 4, light green, 15 keywords).

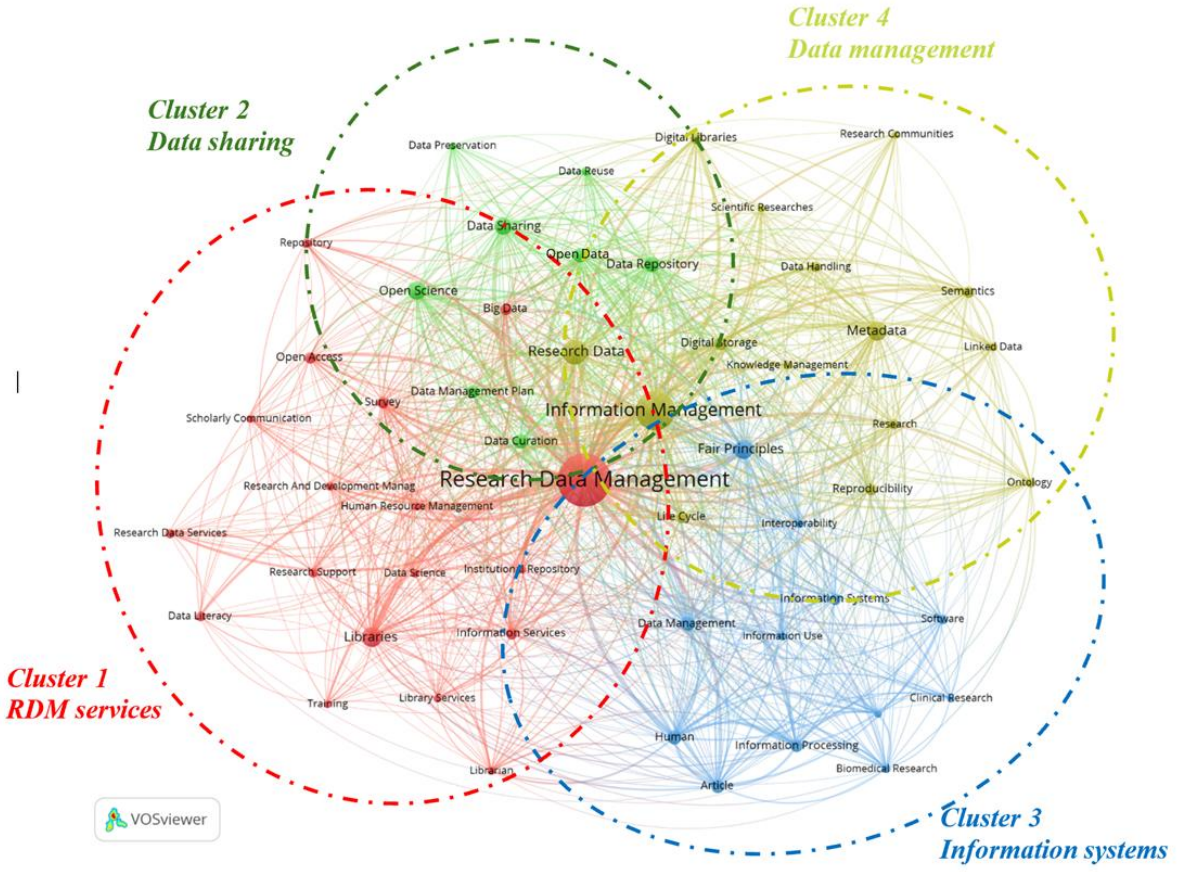


Figure 2. Network Visualization of Keywords Co-occurrence, depicting four distinct clusters - RDM services, Data sharing, Information systems, and Data management.

All keywords, showing the full counting of the occurrences, indicate the total number of times a term appears across all publications (Van Eck & Waltman, 2020). Within each theme, keywords occurring in the respective cluster were sorted first by cluster and then by total link strength. The total link strength refers to the number of publications in which two or more keywords occur together (Van Eck & Waltman, 2020). The primary keyword "Research Data Management" is placed in RDM services theme (Cluster 1). It has the highest occurrence (655, 59.28%) and total link strength (2036, 51.27%) overall and its cluster. The keyword "Open Science" is the most common occurrence (76, 20.60%) with the highest total link strength (329,

Ismail, M. I., Abrizah, A., & Samsuddin, S. F. (2025). Knowledge mapping of research data management: Uncovering themes and developments through co-occurrence and bibliometric analysis. *Record and Library Journal, 11*(1), 234-258. DOI: 10.20473/rlj.V11-I1.2025.234-258.

19.80%) in Data sharing theme (Cluster 2). The highest occurrence in Information systems (Cluster 3) is "FAIR Principles" (82, 20.81%), which has the highest total link strength (363, 16.50%). "Information Management" is the highest occurrence (346, 38.88%) with the highest total link strength (1421, 35.82%) in Data management theme (Cluster 4) and is the second-highest occurrence overall. While "Research Data" is at the third place of highest occurrence and total link strength overall, it has been placed in the Data management (Cluster 4), which was immediately after "Information Management". The findings show that only "Research Data Management" and "Information Management" have more than 300 occurrences with more than 1000 total link strengths, while "Research Data Management" has more than 50% in both occurrences and total link strength for overall keywords in this study. However, these two keywords are not positioned in one cluster. All other keywords have below 150 occurrences. It is worth mentioning that some keywords were low in occurrences but were high in total link strength around the keywords' co-occurrences network. For example, some keywords were between 15 and 20 occurrences, but the total link strength was between 101 and 122, such as "Interoperability" (Cluster 3), "Medical Research" (Cluster 3), "Information Use" (Cluster 3), "Human Resource Management" (Cluster 1), "Librarian" (Cluster 1), "Institutional Repository" (Cluster 1), "Linked Data" (Cluster 4) and "Data Reuse" (Cluster 2). This indicates that even if these keywords occur infrequently, they may have strong connections with other important keywords, indicating their significance within the specific theme or cluster.

RDM Subject Areas

This study classified subject areas based on the publications in each cluster, as presented in Table 2. The subject areas align with the Scopus database, encompassing 26 specific subject areas along with additional general subject areas containing multidisciplinary journals, namely Life Sciences, (ii) Physical Sciences, (iii) Social Sciences, and (iv) Health Sciences (García et al., 2011).

Table 2 illustrates that RDM as a topic spans various subject areas. As certain publications may be associated with more than one subject area, the data in Table 2 reflects a count greater than the actual total number of publications. The theme that has the highest number of publications covering the subject area is RDM services (Cluster 1, 1102), followed by Data sharing (Cluster 2, 746), followed by Data management (Cluster 4, 418), and Information systems (Cluster 3, 353). Computer Science and Social Sciences emerge as the two highly represented subject areas in RDM publications. Computer Science subject area exhibits the highest number of publications in all four clusters, accounting for more than 25% of the total publications: RDM services (Cluster 1, 33.76%), Data sharing (Cluster 2, 42.49%), Information systems (Cluster 3, 27.48%), and Data management (Cluster 4, 35.17%). Social Sciences represents more than 25% of total publications in two clusters: RDM services (Cluster 1, 28.04%) and Data sharing (Cluster 4, 28.71%). It is evident that Computer Science and Social Sciences exert dominance in RDM as a research topic. The total number of publications in these two subject areas significantly surpasses those in other subject areas, where most have fewer than 50 publications in each cluster (except Mathematics).

Table 2. RDM Subject Areas by Cluster

Subject Area	Total Publications (TP)							
	Cluster 1 RDM SERVICES		Cluster 2 DATA SHARING		Cluster 3 INFORMATION SYSTEMS		Cluster 4 DATA MANAGEMENT	
Agricultural and Biological Sciences	10	0.91%	4	0.54%	5	1.42%	3	0.72%

Ismail, M. I., Abrizah, A., & Samsuddin, S. F. (2025). Knowledge mapping of research data management: Uncovering themes and developments through co-occurrence and bibliometric analysis. *Record and Library Journal, 11*(1), 234-258. DOI: 10.20473/rlj.V11-I1.2025.234-258.

Arts and Humanities	30	2.72%	8	1.07%	5	1.42%	11	2.63%
Biochemistry, Genetics and Molecular Biology	23	2.09%	18	2.41%	18	5.10%	7	1.67%
Business, Management and Accounting	25	2.27%	14	1.88%	4	1.13%	9	2.15%
Chemical Engineering	12	1.09%	10	1.34%	7	1.98%	1	0.24%
Chemistry	14	1.27%	13	1.74%	10	2.83%	2	0.48%
Computer Science	*372*	*33.76%*	*317*	*42.49%*	*97*	*27.48%*	*147*	*35.17%*
Decision Sciences	37	3.36%	37	4.96%	13	3.68%	20	4.78%
Dentistry	1	0.09%	0	0.00%	0	0.00%	0	0.00%
Earth and Planetary Sciences	13	1.18%	7	0.94%	4	1.13%	4	0.96%
Economics, Econometrics and Finance	5	0.45%	0	0.00%	0	0.00%	2	0.48%
Energy	3	0.27%	1	0.13%	1	0.28%	2	0.48%
Engineering	49	4.45%	39	5.23%	22	6.23%	21	5.02%
Environmental Science	14	1.27%	11	1.47%	9	2.55%	2	0.48%
Health Professions	20	1.81%	13	1.74%	19	5.38%	6	1.44%
Immunology and Microbiology	3	0.27%	3	0.40%	3	0.85%	0	0.00%
Materials Science	6	0.54%	5	0.67%	2	0.57%	1	0.24%
Mathematics	81	7.35%	81	10.86%	16	4.53%	35	8.37%
Medicine	50	4.54%	29	3.89%	47	13.31%	16	3.83%
Multidisciplinary	4	0.36%	3	0.40%	3	0.85%	3	0.72%
Neuroscience	4	0.36%	1	0.13%	5	1.42%	2	0.48%
Nursing	2	0.18%	2	0.27%	3	0.85%	1	0.24%
Pharmacology, Toxicology and Pharmaceutics	5	0.45%	4	0.54%	4	1.13%	0	0.00%
Physics and Astronomy	9	0.82%	6	0.80%	3	0.85%	2	0.48%
Psychology	1	0.09%	1	0.13%	1	0.28%	1	0.24%
Social Sciences	*309*	*28.04%*	*119*	*15.95%*	*52*	*14.73%*	*120*	*28.71%*
Total	1102	100.00 %	746	100.00 %	353	100.00 %	418	100.00%

Temporal Trend in RDM Publications

RDM publications associated with the keyword "Research Data Management" emerge as the central node in the keywords co-occurrence network (see Figure 2 and 3). Figure 3 shows the average publication year in which a keyword or a term occurs or the average publication year published by a source, an author, an organization, or a country (van Eck & Waltman, 2020). The majority of publications are dated from 2019 onwards, with a particular emphasis on RDM services (Cluster 1) and Data sharing (Cluster 2). These clusters feature numerous keywords highlighted in yellow, orange, and red. Cluster 2, focused on Data Sharing, specifically reveals recent keywords introduced in publications from 2018 onwards (depicted

Ismail, M. I., Abrizah, A., & Samsuddin, S. F. (2025). Knowledge mapping of research data management: Uncovering themes and developments through co-occurrence and bibliometric analysis. *Record and Library Journal, 11*(1), 234-258. DOI: 10.20473/rlj.V11-I1.2025.234-258.

in red). Keywords such as “Data Management Plan,” “Open Science,” “Reproducibility,” “Interoperability,” and “FAIR Principles” are prominent nodes highlighted in red within the Data Sharing (Cluster 2) and Information Systems (Cluster 3) themes in this co-occurrence network. While the term “Libraries” stands out as the largest node in yellow within Cluster 1, corresponding to the theme of RDM services. “Research Data Services” and “Data Literacy” are identified as the most recently published keywords in Cluster 1. Notably, both Cluster 3 (Information Systems) and Cluster 4 (Data management) exhibit a multitude of keywords with average publications dating before 2018.

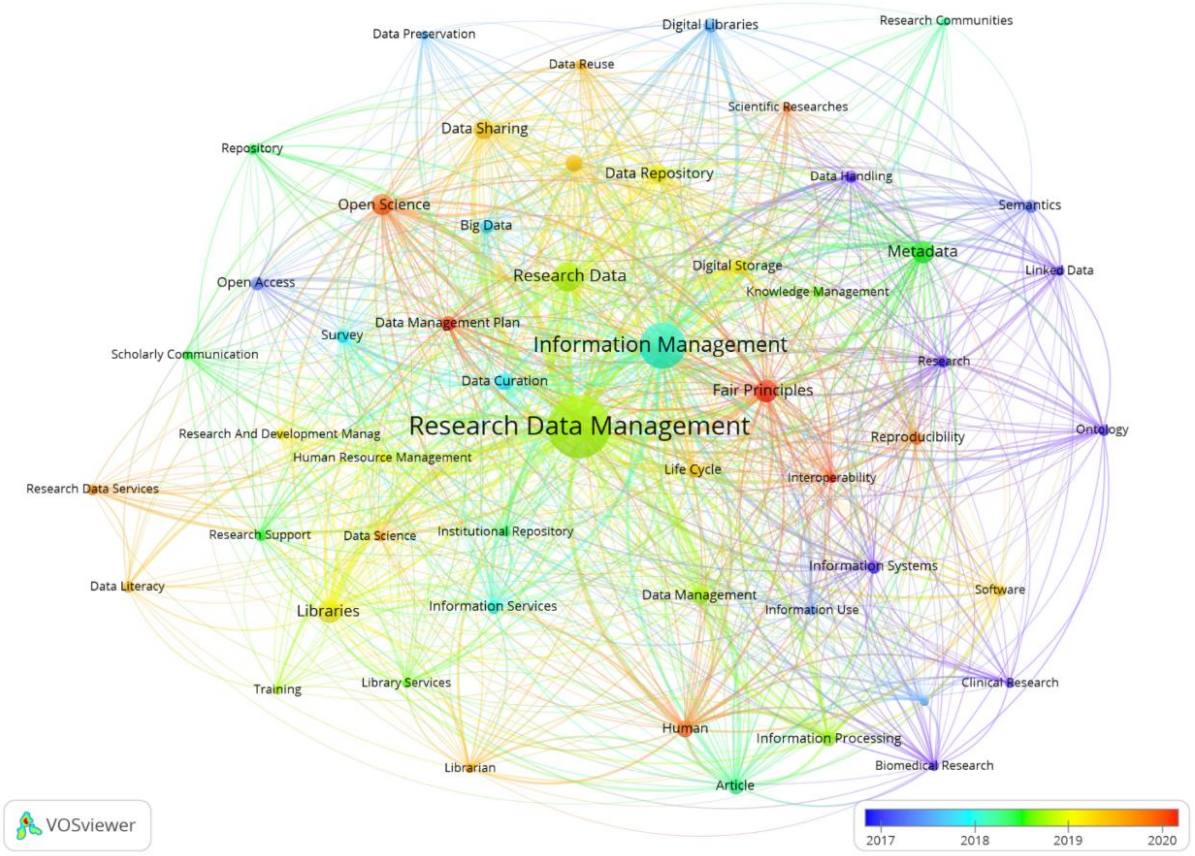


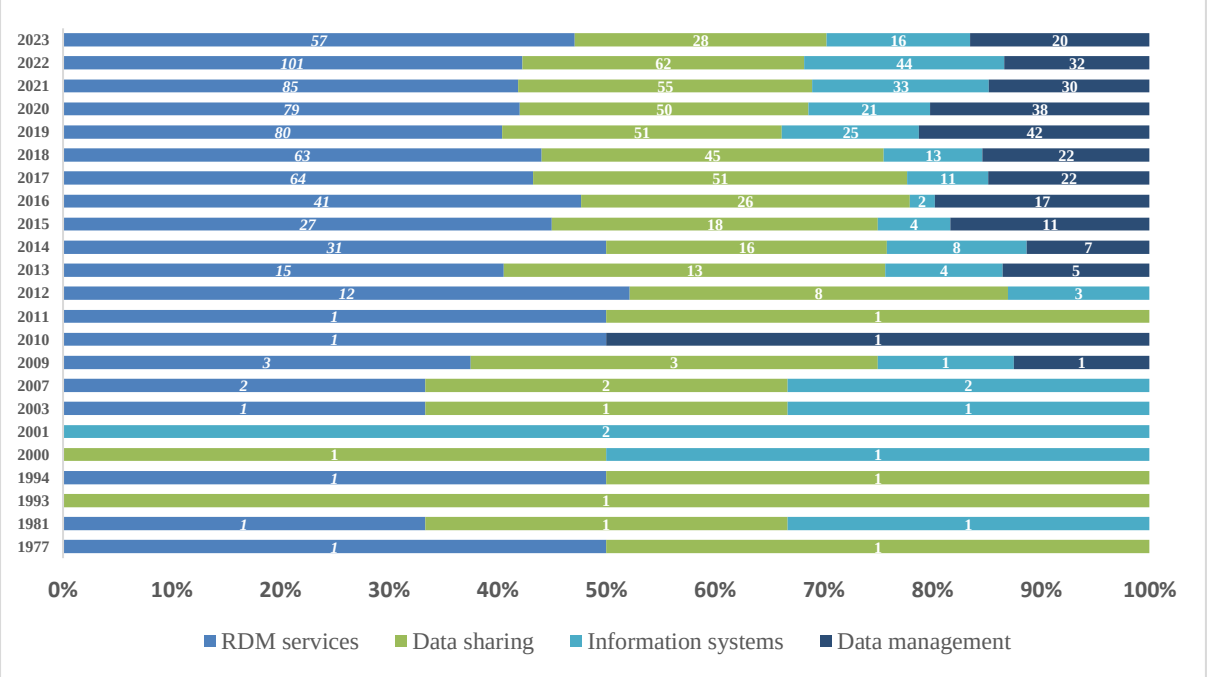
Figure 3. Keywords co-occurrence network by year of publications

Figure 4 illustrates the distribution of RDM publications from 1977 to August 2023 across different research areas/clusters. Each cluster exhibits varying publication counts: RDM services (Cluster 1, 666), Data sharing (Cluster 2, 435), Information systems (Cluster 3, 192), and Data management (Cluster 4, 248). Between 1977 and 2011, each research cluster recorded no more than three publications. However, a significant increase in publications became evident from 2012 onward.

The study indicates that the majority of publications were concentrated within the RDM services cluster, surpassing the other three clusters. RDM services showed a notable increase, going from 12 publications in 2012 to over 80 publications in 2019. Similarly, Data sharing also experienced an increase from 8 publications in 2012 to more than 50 publications in 2017. The growth in the number of Information systems was comparatively modest, progressing from only three publications in 2012 to over 20 in 2019. In contrast, Data management witnessed an increase as well, starting with 5 publications in 2013 and reaching 42 publications in 2019. Overall, each cluster in the study demonstrated that more than 10% of publications were recorded from 2019 onwards, with the exception of Data sharing, which achieved this milestone in 2017.

Ismail, M. I., Abrizah, A., & Samsuddin, S. F. (2025). Knowledge mapping of research data management: Uncovering themes and developments through co-occurrence and bibliometric analysis. *Record and Library Journal, 11*(1), 234-258. DOI: 10.20473/rlj.V11-I1.2025.234-258.

RDM publications demonstrated a gradual and ongoing increase, reflecting a sustained interest in the field of RDM. However, it is important to note that the number of publications for 2023 was incomplete as of August 2023, when the data for this study was collected. These highlights the growing interest in RDM, particularly in areas such as RDM services and Data sharing, which possibly continue to attract significant scholarly attention.



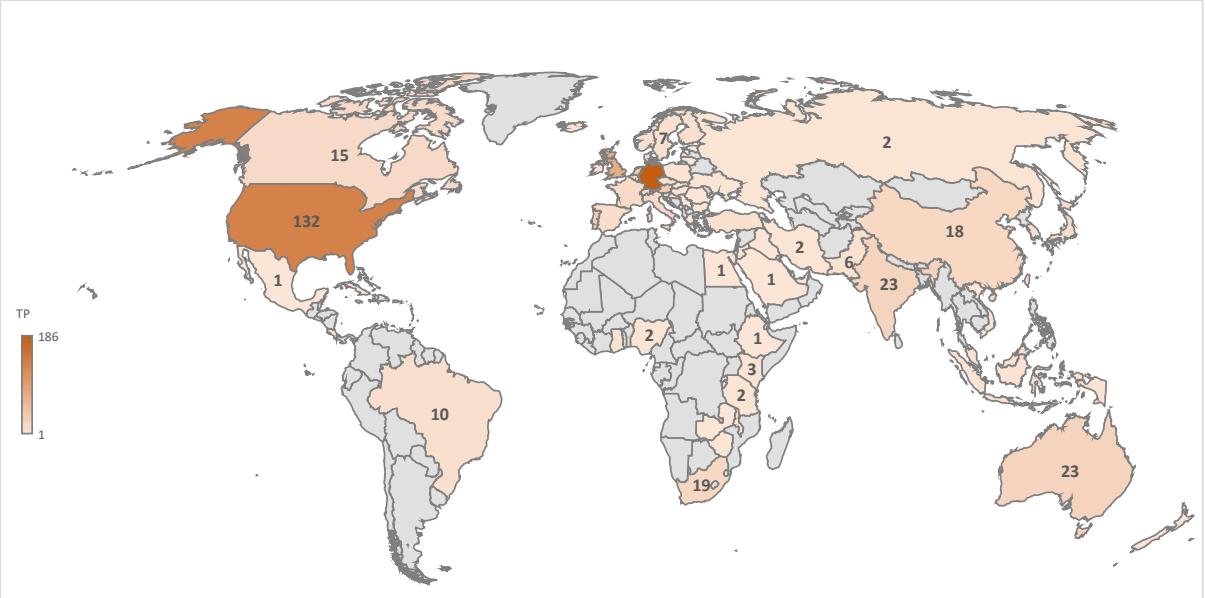
* One publication may represent more than one cluster

Figure 4. Distribution of RDM Publications (1977–August 2023) across Research Clusters

Geographical Distribution of RDM Publications

Understanding the concentration of research on RDM offers a worldwide perspective on the prevalence and significance of these studies, enabling researchers and policymakers to recognize the extensive international engagement in RDM. Therefore, the study meticulously analyzes the 684 RDM publications, revealing contributions from 70 countries and involving 829 authors. Figure 5 presents the global distribution of RDM publications, emphasizing the significant contributions from three key countries - Germany (186, 22.44%), the United States (132, 15.92%), and the United Kingdom (132, 15.92%). The analysis further explores the geographic distribution of RDM publications across continents, highlighting Europe as the primary contributor across all clusters (63.6%), particularly in RDM services (25.3%) and Data sharing (18.1%) and Data management (11.1%) (Table 3). Europe's leadership in RDM research arises from various factors, notably its pioneering advocacy for open science practices (Donner, 2023; Matusiak & Sposito, 2017). The strong commitment to openness and transparency in publicly funded research has fostered a collaborative and productive research environment in the field of RDM. As data-driven research continues to grow in importance, RDM practices have become integral to publicly funded projects (Salazar et al., 2023). These findings underscore the critical role of RDM in advancing global research initiatives and highlight the need for continued investment and collaboration in this area.

Ismail, M. I., Abrizah, A., & Samsuddin, S. F. (2025). Knowledge mapping of research data management: Uncovering themes and developments through co-occurrence and bibliometric analysis. *Record and Library Journal, 11*(1), 234-258. DOI: 10.20473/rlj.V11-I1.2025.234-258.



TP: Total publications
* One publication may represent more than one country.

Figure 5. Geographical Distribution of RDM Publications by Country

Table 3. Geographic Distribution of RDM Publications across Continents

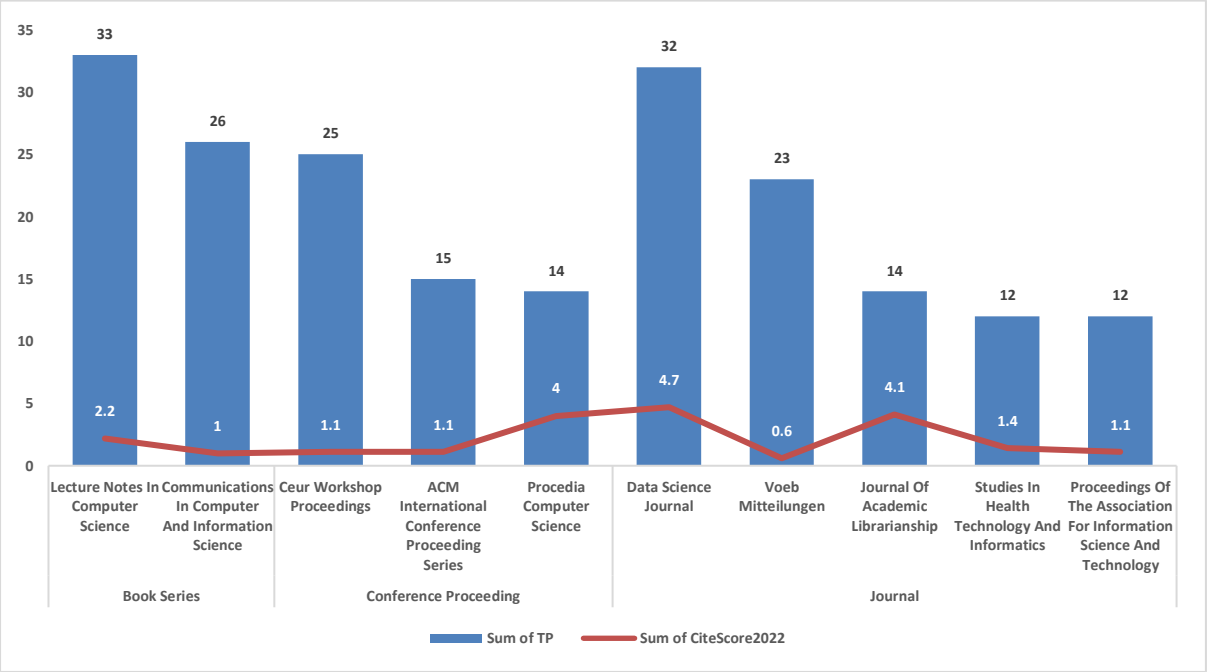
Continent	Total Publications (TP)				
	Cluster 1 RDM SERVICES	Cluster 2 DATA SHARING	Cluster 3 INFORMATION SYSTEMS	Cluster 4 DATA MANAGEMENT	Grand Total
Europe	483 (25.3%)	346 (18.1%)	173 (9.1%)	211 (11.1%)	1213 (63.6%)
North America	144 (7.5%)	90 (4.7%)	57 (3.0%)	41 (2.1%)	332 (17.4%)
Asia	95 (5.0%)	54 (2.8%)	13 (0.7%)	54 (2.8%)	216 (11.3%)
Africa	35 (1.8%)	20 (1.0%)	5 (0.3%)	14 (0.7%)	74 (3.9%)
Oceania	26 (1.4%)	12 (0.6%)	4 (0.2%)	6 (0.3%)	48 (2.5%)
South America	9 (0.5%)	5 (0.3%)	2 (0.1%)	1 (0.1%)	17 (0.9%)
Undefined	5 (0.3%)	1 (0.1%)	1 (0.1%)	1 (0.1%)	8 (0.4%)
Grand Total	797 (41.77%)	528 (27.67%)	255 (13.36%)	328 (17.19%)	1908 (100%)

Document Types and Preferred Source Titles of RDM Publications

The analysis of document types and preferred source titles provides a comprehensive understanding of the RDM research landscape, serving multiple purposes, including facilitating quality assessment and streamlining literature review processes within the field. Of the 684 publications published in 295 source titles, 57% (392) were journals, and 35% (240) were conference papers. Other types included 27 (4%) reviews, 19 (3%) book chapters, three (0.4%) notes, two (0.3%) editorials, and one (0.15%) data paper, representing the varied and diverse sources favoured in RDM publications. Figure 6 presents the top 10 preferred source titles of RDM publications, along with their respective source types and CiteScores. Notably, the top-ranking source in terms of the total number of RDM publications was the book series Lecture Notes in Computer Science (33 publications), followed closely by the journal Data Science Journal (32 publications), and another book series, Communications in Computer and

Ismail, M. I., Abrizah, A., & Samsuddin, S. F. (2025). Knowledge mapping of research data management: Uncovering themes and developments through co-occurrence and bibliometric analysis. *Record and Library Journal, 11*(1), 234-258. DOI: 10.20473/rlj.V11-I1.2025.234-258.

Information Science (26 publications). Additionally, the conference proceedings Ceur Workshop Proceedings (25 publications) and the journal VOEB- Mitteilungen (23 publications) are also notable sources favoured by researchers in the field of RDM. Within the top 10 preferred source titles, four source titles - Data Science Journal, Communications in Computer and Information Science, Procedia Computer Science, and Studies in Health Technology and Informatics - are consistently present across all RDM clusters. The remaining titles are represented in at least one cluster.



* One publication may represent more than one cluster and continent.

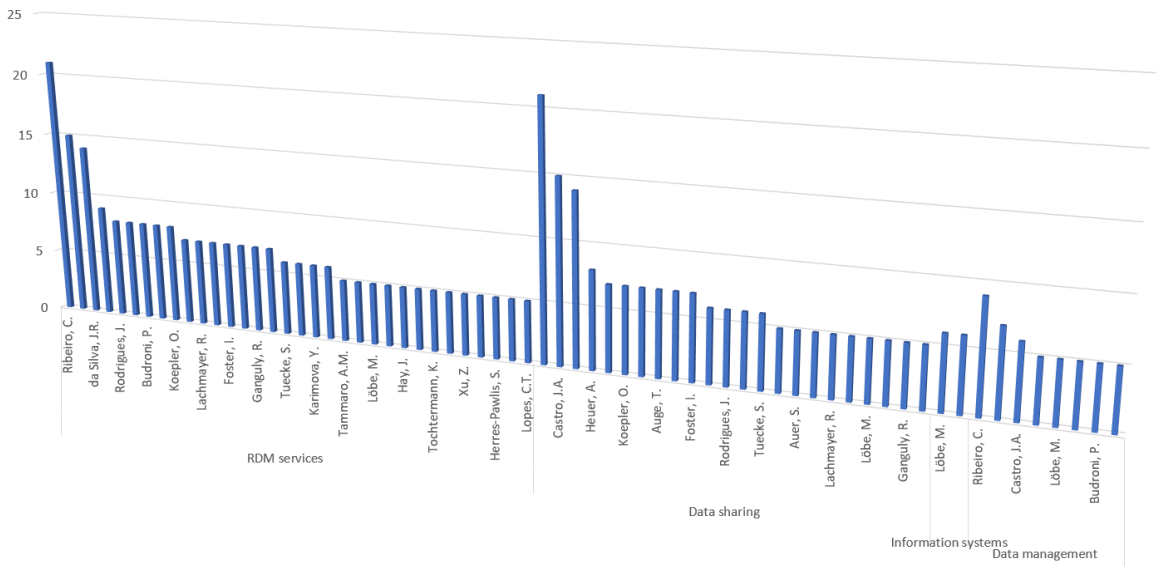
Figure 6. Top 10 Preferred Source Titles of RDM Publications Highlighting Source Types and CiteScores

Authorship

Recognizing prolific authors in an emerging field like RDM is essential for acknowledging thought leaders, fostering collaboration, and providing benchmarks for quality research. It guides new researchers, shapes the field's direction, and aids in resource allocation and research collaboration, contributing to the growth and sustainability of the discipline. Figure 7 presents the most prolific authors in each cluster who have made significant contributions to the field, each having at least five publications in this study. Notably, all 33 authors (out of 1915 authors) with such substantial contributions were affiliated with RDM services. Conversely, there were minimal authors with at least five publications in Information Systems (2 authors out of 919 authors) and Data Management (8 authors out of 947 authors). In contrast, Data Sharing stood out with 23 authors (out of 1755 authors) contributing at least five publications, making it the second most productive cluster after RDM services in terms of prolific authors.

Several authors showed similar publication numbers due to co-authorship and shared affiliations. Examples include Cox, A.M. (9 publications) and Pinfield, S. (6 publications) from the University of Sheffield in RDM services; Ribeiro, C. (21 publications), Castro, J.A. (15 publications), and da Silva, J.R. (14 publications) from Universidade Do Porto in Data Sharing; and Data Management. Additionally, Koepler, O. (5 publications) and Auer, S. (4 publications) from Leibniz Information Centre of Science and Technology University Library in Information Systems (Cluster 3) displayed similar collaborative patterns.

Ismail, M. I., Abrizah, A., & Samsuddin, S. F. (2025). Knowledge mapping of research data management: Uncovering themes and developments through co-occurrence and bibliometric analysis. *Record and Library Journal, 11*(1), 234-258. DOI: 10.20473/rlj.V11-I1.2025.234-258.



* One publication of one author may represent more than one cluster.

Figure 7. Most productive authors in RDM by clusters (at least five publications)

Citation Impact

The study also seeks to identify influential RDM publications. Figure 8 displays the total number of publications and citations for four RDM clusters, offering an overview of the performance and impact of RDM research within each cluster. RDM services received 3775 citations, which is the highest number of citations received among the clusters for this study (Cluster 1, 666 publications). Data sharing received 2224 citations (Cluster 2, 435 publications), and it was placed in the second highest number of citations received, followed by Data management, which received 1533 citations (Cluster 4, 248 publications), and Information systems received 1118 citations (Cluster 3, 192 publications).

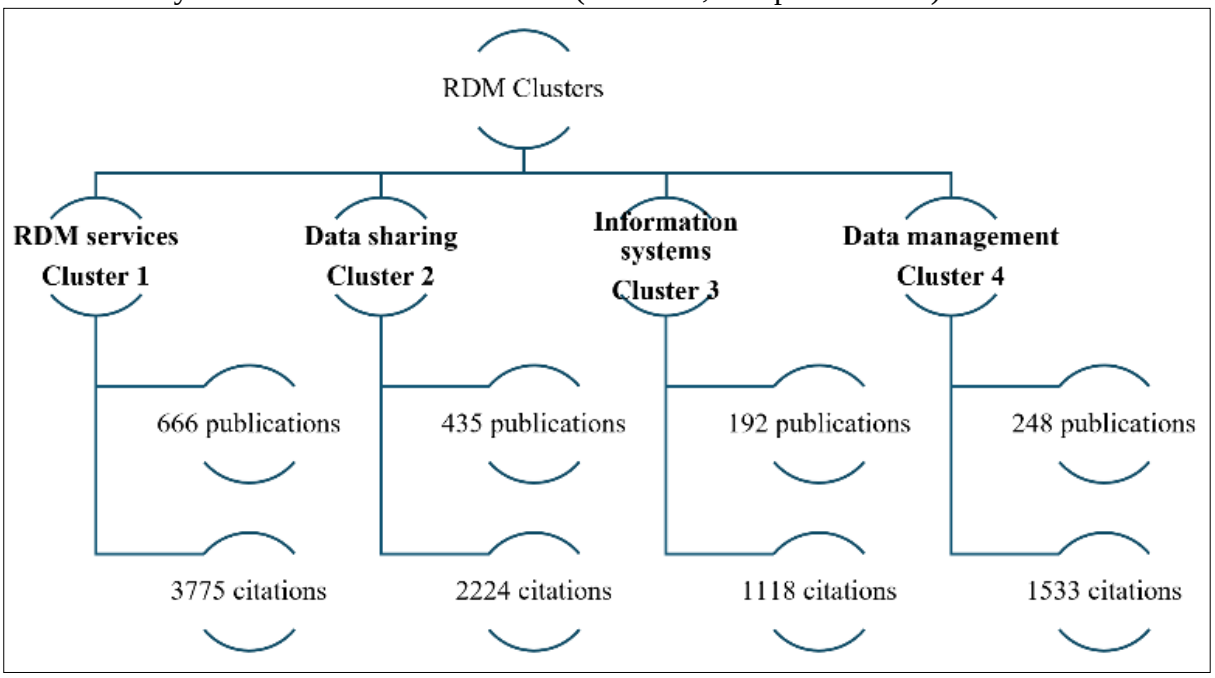


Figure 8. Citation received for RDM Clusters

Figure 9 illustrates the citation mapping for RDM publications with a minimum of 30 citations. The study identified 24 publications meeting this criterion, showcasing the RDM top-

Ismail, M. I., Abrizah, A., & Samsuddin, S. F. (2025). Knowledge mapping of research data management: Uncovering themes and developments through co-occurrence and bibliometric analysis. *Record and Library Journal, 11*(1), 234-258. DOI: 10.20473/rlj.V11-I1.2025.234-258.

cited works and their interrelationships through link strength (Wahid et al., 2020). Notably, Cox & Pinfield (2014), Pinfield et al. (2014), and Cox et al. (2017) emerged as frequently cited, averaging at least 90 citations. This visualization serves to highlight the central publications in the RDM domain and visually depict the inter-connectedness of the researchers' ideas.

While four distinct clusters emerged from this study, certain publications could be associated with more than one cluster due to co-occurrence networks developed by VOSviewer. Overall, this study reveals notable citation patterns among publications by Cox & Pinfield (2014) having 141 citations, 15.67 citations per year in RDM services and Data management; and Cox et al. (2017), having 114 citations and 19 citations per year in RDM services, Data sharing and Information systems. Pinfield et al. (2014), Koltay (2017), and Cox et al. (2019) are also ranked within the top 10 publications with the highest citations across various thematic clusters. This analysis yields valuable insights into the impact of citations and the thematic relevance of the examined publications within the RDM field.

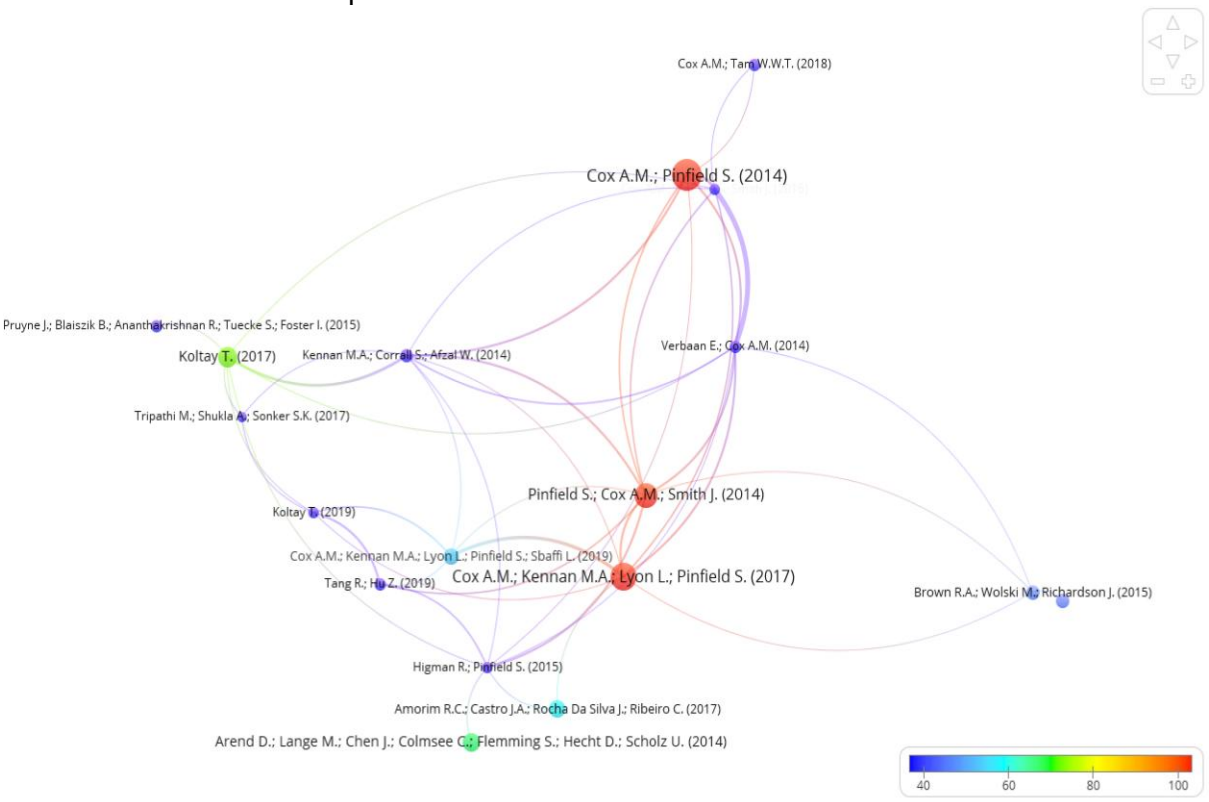


Figure 9. Citation mapping of RDM publications with a minimum of 30 citations, illustrating key works and their interconnections

Discussion

The discussion delves into the outcomes of the preceding co-occurrence and bibliometric analysis, reexamining the study's central research question: "How can the relationships among research data management literature be described and analyzed in a representative, dynamic, and scalable way?" The publication selection from Scopus databases from 1977 to the date of the study in August 2023, centered around the primary keyword "Research Data Management." Utilizing VOSviewer for data analysis, the study uncovers patterns, trends, and key contributors shaping the field, as well as extracted key themes, consolidating them into four clusters.

RDM Themes

Ismail, M. I., Abrizah, A., & Samsuddin, S. F. (2025). Knowledge mapping of research data management: Uncovering themes and developments through co-occurrence and bibliometric analysis. *Record and Library Journal, 11*(1), 234-258. DOI: 10.20473/rlj.V11-I1.2025.234-258.

The ensuing discussion encapsulates the findings within these four clusters, representing the main RDM themes: RDM services, Data sharing, Information systems, and Data management.

RDM Services

RDM research revolves around keywords focused on themes related to RDM services. The co-occurrence analysis highlights that a majority of publications within this theme delve into aspects such as libraries, information services, research support, data literacy, and repositories, indicating a higher frequency of discussion in these areas. Libraries assume a pivotal role in RDM services, leveraging their reliable capabilities and expertise across a spectrum of skills and knowledge (Avuglah & Underwood, 2019). Librarians, equipped with the necessary skills and experience in information services, repositories, and research support, contribute significantly to RDM services. Various library services, including the management of metadata in data repositories and the provision of information services through advocacy and researcher support via training and consultations, are closely linked to RDM services. The scholarly literature consistently underscores the pivotal role of the library as a primary agent in formulating and providing RDM services to researchers (Bunkar & Bhatt, 2020; Chawinga & Zinn, 2020; Cox & Pinfield, 2014; Harrison, 2018; Koltay, 2016; Mushi et al., 2020; Nitecki & Davis, 2019; Pinfield et al., 2014; Tammaro et al., 2019). Efforts to enhance RDM awareness are facilitated through advocacy programs, training sessions, and workshops conducted by libraries (Li et al., 2019; Marlina & Purwandari, 2019; Wiljes & Cimiano, 2019; Wiorogórska et al., 2018).

Data Sharing

The second theme is closely linked to keywords co-occurring in the context of data sharing issues. Notably, the primary keyword "Research Data Management" is positioned within this theme, reflecting its close association with significant terms. As suggested by (Zhu et al., 2015), the prevalence of high-frequency keywords in a particular cluster indicates a focused concentration on a specific topic, particularly emphasizing the practical application of research findings. This theme prominently features the co-occurrence of the primary keyword with other significant terms centered around data sharing, such as Open science, Data sharing, Data curation, and Open data. This association aligns with the findings of Gupta et al. (2021), emphasizing that data sharing ranks among the most frequently co-occurring keywords in RDM publications on the Web of Science. Data sharing, a key component of open science, is highlighted as essential for excellent RDM, serving as a prerequisite for open science and RDM policies (Timmermann, 2019). The study by Vilar & Zabukovec (2019) underscores the growing importance of RDM for researchers and supporting research units within the context of the open science movement. Moreover, open data emerges as a crucial element in achieving the open science agenda, with its frequent mention of data sharing and reuse (Mosconi et al., 2019). The integration of RDM and open science strategies requires comprehensive research support (Rice, 2019). The FAIR principles underscore the importance of making data Findable, Accessible, Interoperable, and Reusable (Schöpfel et al., 2018). Adherence to these principles becomes crucial as researchers are encouraged to participate actively in RDM and embrace openness. The interrelated keywords within this theme facilitate a clear understanding of their relationships.

Information Systems

This theme is closely associated with issues pertaining to information systems in the development and implementation of RDM. Information systems play a crucial role in

Ismail, M. I., Abrizah, A., & Samsuddin, S. F. (2025). Knowledge mapping of research data management: Uncovering themes and developments through co-occurrence and bibliometric analysis. *Record and Library Journal, 11*(1), 234-258. DOI: 10.20473/rlj.V11-I1.2025.234-258.

collecting, storing, and processing data, which, in turn, contributes to the generation of knowledge and digital products. Particularly in clinical and biomedical research, information systems are extensively utilized for managing research data, given the digitization of most data in these fields (Tang et al., 2018). The realm of information systems in RDM encompasses tasks such as data processing, integration, and retrieval (Pinfield et al., 2014). Beyond the requisite skills and knowledge for developing RDM services, technology infrastructure serves as another critical resource allocation for RDM. This includes software and hardware to support researchers in their RDM activities (Aydinoglu et al., 2017). Challenges arise when the lack of ICT infrastructure hinders the provision of facilities for researchers, impacting storage, preservation, and compliance with institutional requirements, funders, and publishers (Kruse & Thestrup, 2014). Exploring future aspects of RDM through research on information systems and technology infrastructure issues promises to enhance the development of activities and services in the realm of RDM.

Data Management

The study reveals a substantial concentration of RDM publications within the realms of computer sciences and social sciences, particularly with a distinct focus on the LIS domain. The keywords within this theme are intricately entwined with the broader field of information management, and are directed towards various data management activities, with a notable focus on metadata management. Effective metadata management is pivotal for the flexibility and efficiency of data management, as metadata can be accessed on data platforms or databases, contributing to the long-term preservation of research data (Finkel et al., 2020). However, a significant challenge arises from the fact that many researchers have not consistently used standardized ontologies or metadata schemas. This lack of standardization poses challenges for libraries in providing and establishing semantically linked sources for research data (Schirrwagen et al., 2019). It is important to recognize that metadata for research data may differ from that of research publications, such as journal articles. The creation of metadata for research data often necessitates direct contributions from researchers in describing its specifics. In contrast, librarians typically play a role in accurately describing the metadata for research publications. In light of these challenges, research on data management should be a focal point for researchers involved in handling data, spanning from creating metadata to data publication throughout the entire research lifecycle. Researchers play a critical role in exploring various facets of RDM to meet both personal needs and the requirements set forth by institutions, funders, and publishers.

RDM literature growth and performance

The bibliometric analysis highlighted a significant increase in RDM publications, particularly beginning in 2012 and continuing to rise. This upward trend suggests a growing interest among researchers in exploring various aspects related to RDM. The present study portrays the sustained expansion of RDM literature, particularly within the theme of "Data Sharing", which encompasses issues pertaining to open science, open data, and data sharing. In line with the findings of Zhang & Eichmann-Kalwara (2019), who identified scientific collaboration, research support services, and data literacy as the top three clusters in RDM literature, the current study corroborates these results by placing research support services and data literacy within the overarching theme of RDM services. Notably, while Zhang & Eichmann-Kalwara (2019) reported an average publication year for data literacy as 2012. The present study indicates that publications on data literacy have experienced average growth from 2017 onward, mirroring trends observed in other themes. Topics related to open science, open data, and data sharing have gained prominence in literature since 2009, which has been marked

Ismail, M. I., Abrizah, A., & Samsuddin, S. F. (2025). Knowledge mapping of research data management: Uncovering themes and developments through co-occurrence and bibliometric analysis. *Record and Library Journal, 11*(1), 234-258. DOI: 10.20473/rlj.V11-I1.2025.234-258.

by the development of various open science initiatives (Zhang et al., 2018). The ongoing discussion and proliferation of research on open data underscores the expanding scope and significance of RDM-related issues in contemporary research.

This study highlights the significant influence of the developed nations, predominantly the Global North, in shaping the RDM research landscape. This leadership stems from the Global North's advanced research infrastructure, substantial funding opportunities, and institutional commitment to open science and transparency in publicly funded research (European Commission, 2019; UNESCO, 2021), enabling the Global North to set benchmarks for RDM practices and policies. The prominence of the Global North in RDM research highlights its pivotal role in driving global advancements, shaping international standards, and influencing the adoption of best practices worldwide. However, it also highlights the need for greater inclusivity and support for RDM initiatives in the Global South to ensure a more equitable and globally representative research ecosystem (CODATA, 2020).

Although the volume of RDM publications in developing nations is relatively small, there is a noticeable upward trend. The analysis of document types and source titles provides a comprehensive view of the RDM research landscape, facilitating quality assessment and streamlining literature reviews. Predominantly, journals and conference papers stand out among various publication types, as also reported in other studies (Ahmi et al., 2018; Al-Lawati et al., 2022; Lam et al., 2022), providing a well-rounded and credible foundation for bibliometric analyses. The emphasis on top source titles within the journal category of Computer Sciences and LIS reflects the core journals publishing RDM, underscoring the significance of these established outlets in shaping and disseminating research in the field. Journals, known for their thorough peer-review processes, contribute to the scholarly discourse, while conference papers provide a platform for the rapid dissemination of emerging trends and novel findings. The combined utilization of these publication types ensures a holistic understanding of the evolving RDM landscape.

This study reveals that "RDM services", which have the highest number of publications, also received the highest number of citations. In contrast, "Information systems" has the fewest publications and the lowest citation count. The correlation between publications and citations can significantly contribute to the perceived value of research (Bourke & Butler, 1996). Notably, this study identifies RDM publications with a minimum of 30 citations, revealing an interesting observation: the most cited publications did not necessarily have the highest total link strength. For instance, the publication with the highest link strength received only 37 citations among those meeting the minimum citation threshold. Interestingly, the top four publications with the highest total link strength were authored by the same individual, potentially suggesting a connection through similar affiliations. As with any bibliometric analysis, including citation and co-citation studies which help identify core literature (Osareh, 1996), this study provides valuable insights into the structure of scientific knowledge and the evolving landscape of RDM (Osareh, 1996).

Conclusion

This study has provided a comprehensive overview of the evolving landscape of RDM, providing researchers, policymakers, and institutions valuable insights into the current state and trends of RDM research. This paper distinguishes itself from previous bibliometric studies on RDM literature by initiating the exploration with the development of clusters, identifying four main themes based on keyword co-occurrence networks. Most RDM publications have predominantly centered around the subject areas of computer science and social science, illustrating their interconnected nature. However, a notable emphasis has been explicitly placed on the domain of LIS. In recent years, there has been a marked increase in publications

Ismail, M. I., Abrizah, A., & Samsuddin, S. F. (2025). Knowledge mapping of research data management: Uncovering themes and developments through co-occurrence and bibliometric analysis. *Record and Library Journal, 11*(1), 234-258. DOI: 10.20473/rlj.V11-I1.2025.234-258.

discussing critical issues such as open science, data sharing, and open data, as well as the expanding role of libraries in RDM services. These findings underscore the dynamic and interdisciplinary nature of RDM research, reflecting its growing relevance in addressing contemporary scholarly and institutional challenges.

By uncovering geographical disparities in RDM research output, the study sheds light on the global distribution of knowledge and highlights opportunities for international collaboration and knowledge exchange. The identification of influential authors and top-cited papers serves as a roadmap for researchers, guiding them towards key contributions and facilitating a deeper understanding of the seminal works shaping the field. In a broader context, the study contributes to the ongoing discourse on open science practices, collaborative research, and the interdisciplinary nature of RDM. The insights derived from this analysis not only inform decision-making and resource allocation but also support the strategic planning of future research initiatives within the rapidly evolving RDM landscape. Furthermore, the findings emphasize the critical role of libraries in strengthening RDM services, enabling them to better support researchers and effectively address their evolving needs and requirements. This study thus serves as a foundation for advancing RDM practices and fostering a more collaborative and inclusive research environment.

The findings of this bibliometric study contribute to the theoretical advancement of RDM by mapping its knowledge domains and identifying potential research directions. Through cluster analysis, the study illustrates how RDM scholarship has been shaped by interdisciplinary influences, including open science, data sharing, information systems, and data management. Furthermore, it underscores the role of theoretical frameworks, such as the FAIR principles, in guiding RDM research. The emerging trends identified in this study encourage scholars to refine existing models or develop novel frameworks to address the evolving challenges in RDM.

It is important to acknowledge the limitations of the study. While the data collected is extensive, it may not encompass the entirety of the dynamic and evolving RDM landscape. Although the Scopus database is one of the largest and most comprehensive sources, it does not cover all journals, particularly those unindexed, which may contain relevant RDM-related scholarly articles that were overlooked (Sweileh et al., 2017). The study's reliance on bibliometric analysis, while valuable, may not fully capture the qualitative nuances of individual contributions and evolving research trends. Additionally, the cutoff date of August 2023 means that recent developments and publications are not incorporated into the findings. These limitations provide opportunities for future research to delve deeper into the evolving landscape of RDM, considering both quantitative and qualitative dimensions. Despite these constraints, the study serves as a foundational exploration, offering valuable insights into the multifaceted dynamics of RDM. This study also advocates for further exploration and discussion of RDM and its associated issues. Given that RDM is still in the developmental and implementation stages in many countries, future research has the potential to capture the attention of researchers and practitioners, contributing to the growing knowledge base of RDM and its significance in the global research ecosystem.

Acknowledgments

We would like to acknowledge and thank all those who have given valuable contributions to this study

Authors' Contributions

All authors have contributed to the final manuscript. The contribution of all authors: conceptualization, methodology, formal analysis, writing original draft preparation, writing

Ismail, M. I., Abrizah, A., & Samsuddin, S. F. (2025). Knowledge mapping of research data management: Uncovering themes and developments through co-occurrence and bibliometric analysis. *Record and Library Journal, 11*(1), 234-258. DOI: 10.20473/rlj.V11-I1.2025.234-258.

review, and editing. All authors have read and agreed to the published version of the manuscript.

Conflict of Interest

All authors have no conflict of interest related to this study.

Funding

The author(s) received no financial support for the publication of this article

References

Ahmi, A., Rahim, S. A., & Elbardan, H. (2018). A global trend of the electronic supply chain management (e-SCM) research: A bibliometric analysis. *International Journal of Supply Chain Management, 7*(5), 535–542.

Al-Lawati, E. H., Abdul Kohar, U. H., & Shahrin Suleiman, E. (2022). Entrepreneurial culture in educational institutions: A scoping review. *Cogent Business and Management, 9*(1). https://doi.org/10.1080/23311975.2021.1997237

Andrikopoulou, A., Rowley, J., & Walton, G. (2022). Research Data Management (RDM) and the evolving identity of academic libraries and librarians: A literature review. *New Review of Academic Librarianship*, *28*(4), 349–365. https://doi.org/10.1080/13614533.2021.1964549

Ashiq, M., Usmani, M. H., & Naeem, M. (2022). A systematic literature review on research data management practices and services. *Global Knowledge, Memory and Communication, 71*(8–9), 649–671. https://doi.org/10.1108/GKMC-07-2020-0103

Avuglah, B. K., & Underwood, P. G. (2019). Research data management (RDM) capabilities at the University of Ghana, Legon. *Library Philosophy and Practice*, *2019*. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85066066748&partnerID=40&md5=52b10dfefaa0f4cef9146b7930947bea

Aydinoglu, A. U., Dogan, G., & Taskin, Z. (2017). Research data management in Turkey: perceptions and practices. *Library Hi Tech*, *35*(2), 271–289. https://doi.org/10.1108/LHT-11-2016-0134

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*, *1*(1), 377–386. https://doi.org/10.1162/qss_a_00019

Bourke, P., & Butler, L. (1996). Publication types, citation rates and evaluation. *Scientometrics*, *37*(3), 473–494. https://doi.org/10.1007/BF02019259

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 and Information Technology*, *40*(3), 139–146. https://doi.org/10.14429/djlit.40.3.15302

Chawinga, W. D., & Zinn, S. (2020). Research data management at an African medical university: Implications for academic librarianship. *Journal of Academic Librarianship*, *46*(4), 102161. https://doi.org/10.1016/j.acalib.2020.102161

CODATA. (2020). *Open Science for a Global Transformation*. https://doi.org/10.5281/zenodo.3935461

Cox, A. M., & Pinfield, S. (2014). Research data management and libraries: Current activities and future priorities. *Journal of Librarianship and Information Science*, *46*(4), 299–316. https://doi.org/10.1177/0961000613492542

Cruz, M., Dintzner, N., Dunning, A., van der Kuil, A., Plomp, E., Teperek, M., Turkyilmaz-van der Velden, Y., & Versteeg, A. (2019). Policy Needs to Go Hand in Hand with

Ismail, M. I., Abrizah, A., & Samsuddin, S. F. (2025). Knowledge mapping of research data management: Uncovering themes and developments through co-occurrence and bibliometric analysis. *Record and Library Journal, 11*(1), 234-258. DOI: 10.20473/rlj.V11-I1.2025.234-258.

Practice: The Learning and Listening Approach to Data Management. *Data Science Journal*, *18*(1). https://doi.org/10.5334/dsj-2019-045

Donner, E. K. (2023). Research data management systems and the organization of universities and research institutes: A systematic literature review. *Journal of Librarianship and Information Science*, *55*(2), 261–281. https://doi.org/10.1177/09610006211070282

Fadhli, R., Rahmat, B., Ibrahim, C., & Eka Kurnia, B. (2023). Uncovering patterns of author productivity in educational leadership and management in Indonesia. *Record and Library Journal*, *9*(2), 306–318. https://doi.org/10.20473/rlj.V9-I2.2023.306-318

Finkel, M., Baur, A., Weber, T. K. D., Osenbrück, K., Rügner, H., Leven, C., Schwientek, M., Schlögl, J., Hahn, U., Streck, T., Cirpka, O. A., Walter, T., & Grathwohl, P. (2020). Managing collaborative research data for integrated, interdisciplinary environmental research. *Earth Science Informatics*, *13*(3), 641–654. https://doi.org/10.1007/s12145-020-00441-0

Flores, J. R., Brodeur, J. J., Daniels, M. G., Nicholls, N., & Turnator, E. (2015). Libraries and the research data management landscape. In *The Process of Discovery: The CLIR Postdoctoral Fellowship Program and the Future of the Academy* (pp. 82–102). http://www.clir.org/pubs/reports/pub167/

García, J. A., Rodriguez-Sánchez, R., & Fdez-Valdivia, J. (2011). Ranking of the subject areas of Scopus. *Journal of the American Society for Information Science and Technology*, *62*(10), 2013–2023. https://doi.org/10.1002/asi.21589

Gupta, N., Arora, S., & Chakravarty, P. R. (2021). Science mapping and visualization of Research Data Management (RDM): Bibliometric and scientometric study. *Library Philosophy and Practice*, *2021*(September), 1–23.

Harrison, R. (2018). The academic library and the research office: Providing scholarly communications support at Imperial College London-A Case Study. In *Collaboration and the Academic Library: Internal and External, Local and Regional, National and International* (pp. 143–150). Elsevier. https://doi.org/10.1016/B978-0-08-102084-5.00013-4

Hazmi, N. R., Abrizah, A., & Yanti Idaya, A. M. K. (2023). Research data governance activities for implementation in Malaysia research performing organizations: insights from data practitioners via Delphi study. *Malaysian Journal of Library & Information Science*, *28*(3), 37–60. https://doi.org/10.22452/mjlis.vol28no3.3

Ismail, M. I., Abrizah, A., & Samsuddin, S. F. (2021). Mapping the knowledge domains of research data management: A co-occurrence analysis. *Reimagining Libraries for a Post-Pandemic World: Proceedings of the International 8th Conference on Libraries, Information and Society, ICoLIS 2021*. http://commonrepo.um.edu.my/12794/

Ismail, M. I., Samsuddin, S. F., & Abrizah, A. (2024). Fostering professional growth in academic librarians: A SoTL-based lens for Research Data Management capacity building. *Malaysian Journal of Library & Information Science*, *29*(3), 171–197. https://doi.org/https://doi.org/10.22452/mjlis.vol29no3.8

European Commission. (2019). *European Open Science Cloud (EOSC) strategic implementation plan* (Sarah. Jones & J.-F. Abramatic, Eds.). Publications Office of the European Union. https://data.europa.eu/doi/10.2777/202370

Koltay, T. (2016). Facing the challenge of data-intensive research: Research data services and data literacy in academic libraries. In W. Baker, D and Evans (Ed.), *Innovation In Libraries and Information Services* (Vol. 35, pp. 45–61). Emerald Group Publishing Limited. https://doi.org/10.1108/S0732-067120160000035008

Kruse, F., & Thestrup, J. B. (2014). Research libraries' new role in research data management,

Ismail, M. I., Abrizah, A., & Samsuddin, S. F. (2025). Knowledge mapping of research data management: Uncovering themes and developments through co-occurrence and bibliometric analysis. *Record and Library Journal, 11*(1), 234-258. DOI: 10.20473/rlj.V11-I1.2025.234-258.

current trends and visions in Denmark. *LIBER Quarterly: The Journal of the Association of European Research Libraries*, *23*(4), 310–335. https://doi.org/10.18352/lq.9173

Lam, W. H., Lam, W. S., Jaaman, S. H., & Lee, P. F. (2022). Bibliometric analysis of information theoretic studies. *Entropy*, *24*(10), 1–13. https://doi.org/10.3390/e24101359

Lee, J. Y., Syn, S. Y., & Kim, S. (2024). Global research trends in research data management: A bibliometrics approach. *Journal of Librarianship and Information Science*. https://doi.org/10.1177/09610006241239083

Li, X., Wu, P., Shen, G. Q., Wang, X., & Teng, Y. (2017). Mapping the knowledge domains of Building Information Modeling (BIM): A bibliometric approach. *Automation in Construction*, *84*(September), 195–206. https://doi.org/10.1016/j.autcon.2017.09.011

Li, Y., Dressel, W., & Hersey, D. (2019). Research data management: What can librarians really help? *20th International Conference on Grey Literature: Research Data Fuels and Sustains Grey Literature, GL 2018*, *2019-Decem*(1), 67–74. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85062960063&partnerID=40&md5=843ffb6d2b81174ab2cc46b010199f5a

Liu, G., Zotoo, I. K., & Su, W. (2020). Research data management policies in USA, UK 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, L., & Mei, S. (2016). Visualizing the GVC research: a co-occurrence network based bibliometric analysis. *Scientometrics*, *109*(2), 953–977. https://doi.org/10.1007/s11192-016-2100-5

Lozano, S., Calzada-Infante, L., Adenso-Díaz, B., & García, S. (2019). Complex network analysis of keywords co-occurrence in the recent efficiency analysis literature. *Scientometrics*, *120*(2), 609–629. https://doi.org/10.1007/s11192-019-03132-w

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

Martín-Martín, A., Thelwall, M., Orduna-Malea, E., & Delgado López-Cózar, E. (2021). Google Scholar, Microsoft Academic, Scopus, Dimensions, Web of Science, and OpenCitations' COCI: A multidisciplinary comparison of coverage via citations. *Scientometrics*, *126*(1), 871–906. https://doi.org/10.1007/s11192-020-03690-4

Masalimova, A. R., Erdyneeva, K. G., Kryukova, N. I., Khlusyanov, O. V., Chudnovskiy, A. D., & Dobrokhotov, D. A. (2023). Bibliometric analysis of augmented reality in education and social science. *Online Journal of Communication and Media Technologies*, *13*(2), e202316. https://doi.org/10.30935/ojcmt/13018

Matusiak, K. K. K. K. K., & Sposito, F. A. (2017). Types of research data management services: An international perspective. *Proceedings of the Association for Information Science and Technology*, *54*(1), 754–756. https://doi.org/10.1002/pra2.2017.14505401144

Mosconi, G., Li, Q., Randall, D., Karasti, H., Tolmie, P., Barutzky, J., Korn, M., & Pipek, V. (2019). Three gaps in opening science. *Computer Supported Cooperative Work (CSCW)*, *28*(3–4), 749–789. https://doi.org/10.1007/s10606-019-09354-z

Mosha, N. F., & Ngulube, P. (2023). Teaching research data management courses in higher learning institutions in Tanzania. *Library Management*, *44*(1/2), 166–179. https://doi.org/10.1108/LM-04-2022-0033

Mushi, G. E., Pienaar, H., & van Deventer, M. (2020). Identifying and implementing relevant

Ismail, M. I., Abrizah, A., & Samsuddin, S. F. (2025). Knowledge mapping of research data management: Uncovering themes and developments through co-occurrence and bibliometric analysis. *Record and Library Journal*, *11*(1), 234-258. DOI: 10.20473/rlj.V11-I1.2025.234-258.

research data management services for the library at the University of Dodoma, Tanzania. *Data Science Journal*, *19*(1). https://doi.org/10.5334/dsj-2020-001

Naseema, S., & Sevukan, R. (2022). Global research trends in research data management (RDM) – A Scientometric view. *International Journal of Information Science and Management*, *20*(4), 117–135. https://doi.org/10.22034/ijism.2022.698421

Nitecki, D. A., & Davis, M. E. K. (2019). Expanding academic librarians' roles in the research life cycle. *Libri*, *69*(2), 117–125. https://doi.org/10.1515/libri-2018-0066

Osareh, F. (1996). Bibliometrics, citation analysis and co-citation analysis: A review of literature II. *Libri*, *46*(4), 217–225. https://doi.org/10.1515/libr.1996.46.4.217

Perrier, L., Blondal, E., Ayala, A. P., Dearborn, D., Kenny, T., Lightfoot, D., Reka, R., Thuna, M., Trimble, L., & MacDonald, H. (2017). Research data management in academic institutions: A scoping review. *PLoS ONE*, *12*(5). https://doi.org/10.1371/journal.pone.0178261

Pinfield, S., Cox, A. M., & Smith, J. (2014). Research data management and libraries: Relationships, activities, drivers and influences. *PLoS ONE*, *9*(12), e114734. https://doi.org/10.1371/journal.pone.0114734

Preeti, & Amit Kumar. (2023). A bibliometric review of publications on research data management. *Library Waves*, *9*(1), 35–46. http://www.librarywaves.com/index.php/lw/article/view/149

Rice, R. (2019). Supporting research data management and open science in academic libraries: A data librarian's view. *Mitteilungen Der Vereinigung Österreichischer Bibliothekarinnen Und Bibliothekare*, *72*(2), 263–273. https://doi.org/10.31263/voebm.v72i2.3303

Schirrwagen, J., Cimiano, P., Ayer, V., Pietsch, C., Wiljes, C., Vompras, J., & Pieper, D. (2019). Expanding the research data management service portfolio at Bielefeld University according to the three-pillar principle towards data FAIRness. *Data Science Journal*, *18*(1), 1–10. https://doi.org/10.5334/dsj-2019-006

Schöpfel, J., Ferrant, C., André, F., & Fabre, R. (2018). Research data management in the French National Research Center (CNRS). *Data Technologies and Applications*, *52*(2), 248–265. https://doi.org/10.1108/DTA-01-2017-0005

Shah, N. U., Ali, N., Hamid, A., & Khan, M. A. (2024). Research visualization trends in research data management (RDM): A bibliometric analysis. *Global Knowledge, Memory and Communication*. https://doi.org/10.1108/GKMC-09-2023-0349

Shiffrin, R. M., Bö, K., & Borner, K. (2004). Mapping knowledge domains. *Proceedings of the National Academy of Sciences*, *101*(Supplement 1), 5183–5185. https://doi.org/10.1073/pnas.0307852100

Stopar, K., & Bartol, T. (2019). Digital competences, computer skills and information literacy in secondary education: mapping and visualization of trends and concepts. *Scientometrics*, *118*(2), 479–498. https://doi.org/10.1007/s11192-018-2990-5

Sweileh, W. M., Al-Jabi, S. W., AbuTaha, A. S., Zyoud, S. H., Anayah, F. M. A., & Sawalha, A. F. (2017). Bibliometric analysis of worldwide scientific literature in mobile - health: 2006-2016. *BMC Medical Informatics and Decision Making*, *17*(1), 1–12. https://doi.org/10.1186/s12911-017-0476-7

Tammaro, A. M., Matusiak, K. K., Sposito, F. A., & Casarosa, V. (2019). Data curator's roles and responsibilities: An international perspective. *Libri*, *69*(2), 89–104. https://doi.org/10.1515/libri-2018-0090

Tang, H., Gao, K., Zhou, Y., Zheng, W., Yang, X., & Zhao, J. (2018). Adopting data analysis and visualization technology to construct clinical research data management and analysis system. *ACM International Conference Proceeding Series*, 49–53.

Ismail, M. I., Abrizah, A., & Samsuddin, S. F. (2025). Knowledge mapping of research data management: Uncovering themes and developments through co-occurrence and bibliometric analysis. *Record and Library Journal, 11*(1), 234-258. DOI: 10.20473/rlj.V11-I1.2025.234-258.

https://doi.org/10.1145/3301761.3301767

Timmermann, M. (2019). A collective challenge: Open Science from the perspective of Science Europe. *Mitteilungen Der Vereinigung Österreichischer Bibliothekarinnen Und Bibliothekare*, *72*(2), 424–430. https://doi.org/10.31263/voebm.v72i2.2831

UNESCO. (2021). *Recommendation on Open Science*. https://doi.org/10.54677/MNMH8546

Us, Y., & Gerulaitienė, N. (2022). Mapping the global research landscape on the nurse resilience management: A bibliometric analysis. *Economics and Business*, *36*(1), 179–198. https://doi.org/10.2478/eb-2022-0012

Van Eck, N. J., & Waltman, L. (2020). VOSviewer manual. In *Leiden: Univeristeit Leiden* (Issue November). University Leiden.

van Eck, N. J., & Waltman, L. (2020). VOSviewer Manual version 1.6.16. *Univeristeit Leiden*, *November*, 1–52. https://www.vosviewer.com/download/f-33t2.pdf

van Eck, N. J., Waltman, L., Eck, N. J., & Waltman, L. (2017). Citation-based clustering of publications using CitNetExplorer and VOSviewer. *Scientometrics*, *111*(2), 1053–1070. https://doi.org/10.1007/s11192-017-2300-7

Verma, S., & Gustafsson, A. (2020). Investigating the emerging COVID-19 research trends in the field of business and management: A bibliometric analysis approach. *Journal of Business Research*, *118*, 253–261. https://doi.org/10.1016/j.jbusres.2020.06.057

Vilar, P., & Zabukovec, V. (2019). Research data management and research data literacy in Slovenian science. *Journal of Documentation*, *75*(1), 24–43. https://doi.org/10.1108/JD-03-2018-0042

Wahid, R., Ahmi, A., & Alam, A. S. A. F. (2020). Growth and collaboration in massive open online courses: A bibliometric analysis. *International Review of Research in Open and Distance Learning*, *21*(4), 292–322. https://doi.org/10.19173/IRRODL.V21I4.4693

Wiljes, C., & Cimiano, P. (2019). Teaching research data management for students. *Data Science Journal*, *18*(1), 1–9. https://doi.org/10.5334/dsj-2019-038

Wilms, K., Stieglitz, S., Buchholz, A., Vogl, R., & Rudolph, D. (2018). Do researchers dream of research data management? *Proceedings of the 51st Hawaii International Conference on System Sciences*. https://doi.org/10.24251/HICSS.2018.556

Wiorogórska, Z., Leśniewski, J., & Rozkosz, E. (2018). Data literacy and research data management in two top universities in Poland. Raising awareness. In *5th European Conference on Information Literacy in the Workplace, ECIL 2017* (Vol. 810, pp. 205–214). Springer Verlag. https://doi.org/10.1007/978-3-319-74334-9_22

Xu, Z. (2022). Research Data Management Training in Academic Libraries: A Scoping Review. *Journal of Librarianship and Scholarly Communication*, *10*(1). https://doi.org/10.31274/jlsc.13700

Zhang, L., & 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*(1). https://doi.org/10.7710/2162-3309.2266

Zhang, Y., Hua, W., & Yuan, S. (2018). Mapping the scientific research on open data: A bibliometric review. *Learned Publishing*, *31*(2), 95–106. https://doi.org/10.1002/leap.1110

Zhu, L., Liu, X., He, S., Shi, J., & Pang, M. (2015). Keywords co-occurrence mapping knowledge domain research base on the theory of big Data in oil and gas industry. *Scientometrics*, *105*(1), 249–260. https://doi.org/10.1007/s11192-015-1658-7

Zhu, X., & Zhang, Y. (2020). Co-word analysis method based on meta-path of subject knowledge network. *Scientometrics*, *123*(2), 753–766.

Ismail, M. I., Abrizah, A., & Samsuddin, S. F. (2025). Knowledge mapping of research data management: Uncovering themes and developments through co-occurrence and bibliometric analysis. *Record and Library Journal*, *11*(1), 234-258. DOI: 10.20473/rlj.V11-I1.2025.234-258.

https://doi.org/10.1007/S11192-020-03400-0

Zotoo, I. K., & Liu, G. (2019). Research data management (RDM) strategy for academic libraries in Ghana: Setting a national development agenda. *OALib*, *06*(04), 1–24. https://doi.org/10.4236/oalib.1105383

Ismail, M. I., Abrizah, A., & Samsuddin, S. F. (2025). Knowledge mapping of research data management: Uncovering themes and developments through co-occurrence and bibliometric analysis. *Record and Library Journal, 11*(1), 234-258. DOI: 10.20473/rlj.V11-I1.2025.234-258.