DOI: 10.21088/ijlis.0973.9548.19125.1

Indian Journal of Library and Information Science
Volume 19 Number 1, January-April 2025

ORIGINAL ARTICLE

The Role of Research Data Management in Accelerating Scientific Discovery

Chaithra A.M.[1], B. Ramesha[2]

HOW TO CITE THIS ARTICLE:

Chaithra A.M., B. Ramesha. The Role of Research Data Management in Accelerating Scientific Discovery. Ind J Lib Inf Sci 2025; 19(1): 07-17.

ABSTRACT

Research data management plays a crucial role in advancing scientific discovery by ensuring the availability, integrity, and accessibility of research data. This article explores the significance of effective research data management practices in accelerating the pace of scientific discovery. It highlights the challenges faced in managing research data and examines the various strategies, tools, and technologies that can be employed to overcome these challenges. The article also discusses the benefits of adopting a proactive approach to research data management, including improved collaboration, data sharing, reproducibility, and the potential for discoveries. Furthermore, it explores emerging trends and future directions in research data management, such as data integration, artificial intelligence, and data-driven discovery. The article concludes by emphasizing the importance of establishing institutional policies and support structures to promote effective research data management practices and foster a culture of data sharing and collaboration.

KEYWORDS

• RDM • Data privacy • Data Sharing • Fair

INTRODUCTION

Scientific progress relies on open access, accountability and retrieval of research data and requires effective research data management (RDM) to drive knowledge generation. This article examines the role, challenges, and solutions of RDM. Managing

AUTHOR'S AFFILIATION:

[1]Research Scholar, Department of Library and Information Science, Bangalore University, Jananabharathi Campus, Bengaluru 560056, India.

[2] President, Association of Teachers Library & Information Science, Department of Library and Information Science, Bangalore University, Jananabharathi Campus, Bengaluru 560056, India.

CORRESPONDING AUTHOR:

Ramesha, President, Association of Teachers Library & Information Science, Department of Library and Information Science, Bangalore University, Jananabharathi Campus, Bengaluru 560056, India.

E-mail: bbramesha@gmail.com

➢ **Received:** 16-12-2025 ➢ **Accepted:** 05-03-2025

Creative Commons Non Commercial CC BY-NC: This article is distributed under the terms of the Creative Commons Attribution NonCommercial 4.0 License (http://www.creativecommons.org/licenses/by-nc/4.0/) which permits non-Commercial use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the Red Flower Publication and Open Access pages (https://rfppl.co.in)

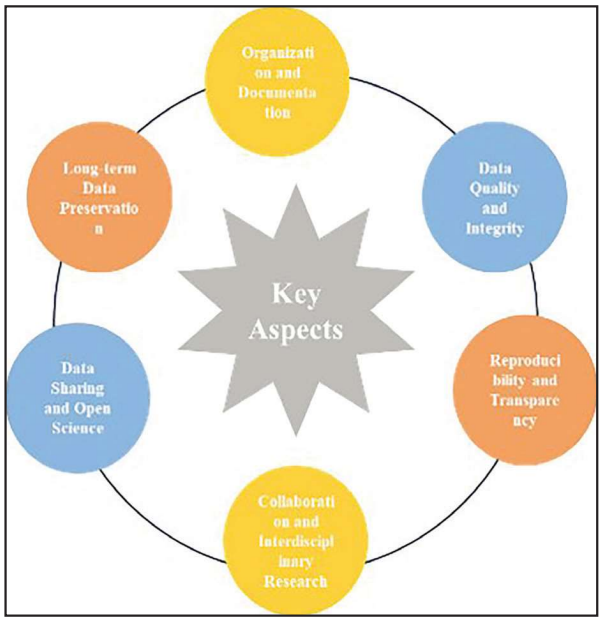
the growing volume, variety, and velocity of data is daunting (Borgman, 2012), yet critical to ensuring data quality and trust (Peng, 2011; Hasselbring & Kao, 2013). Issues such as ownership hinder collaboration and data sharing (Tenopir *et al*., 2015), while reproducibility and transparency require solid documentation (Ioannidis *et al*., 2009). Efficient RDM overcomes these challenges and ensures data volume, quality, security, collaboration and reproducibility. By using effective RDM techniques, researchers can leverage data for scientific breakthroughs and advances.

Research Data Management: Definition and Importance

Research data management (RDM) is about systematically collecting, organizing, storing, documenting, preserving and sharing data throughout the research process to ensure its availability, integrity and accessibility (Klenk *et al*., 2024). Key aspects of RDM include:

1. Organization and documentation: Efficient data organization and documentation saves time, reduces errors and facilitates integration (Moher *et al*., 2020).
2. Data quality and integrity: RDM maintains data quality through rigorous collection and validation processes, increasing reliability (Peng, 2011).
3. Reproducibility and transparency: Transparent, reproducible research is supported by well-documented datasets, promoting scientific accuracy (Singh, Bharti, & Madalli, 2022)
4. Collaboration and interdisciplinary research: RDM facilitates collaboration by enabling data exchange between disciplines and promoting knowledge sharing (Cheung, Van Velden, Lagerburg, & Minderman, 2015)(Gorgolewski & Poldrack, 2016)
5. Data sharing and open science: Open data sharing promotes transparency, innovation and collaboration and enriches the scientific knowledge base (Kratz & Strasser, 2015).
6. Long-term data retention: Retention of data ensures its accessibility for future use and contributes to the accumulation of scientific knowledge (Mayernik *et al*., 2013).

Adopting effective RDM practices enables researchers to address data volume, diversity, quality, security, and sharing challenges, improving research efficiency, reproducibility, and impact, thereby advancing scientific discovery and knowledge.



Objectives of the Study

The aim is to provide practical recommendations and guidelines for improving data management processes. In summary, the study has two main objectives:

1. To analyse current practices and challenges in research data management across various scientific disciplines, particularly their impact on scientific discovery.
2. To propose strategies and guidelines for enhancing scientific discovery through effective research data management.

CURRENT PRACTICES AND CHALLENGES IN RDM ACROSS SCIENTIFIC DISCIPLINES

Research data management practices vary widely across scientific disciplines, reflecting the unique characteristics of data in each field. Understanding these practices and the challenges they pose is critical to maximizing the impact of research data on scientific progress (Borgman, 2012; Tenopir *et al*., 2015).

Data Volume, Variety, and Velocity

Managing research data presents significant challenges due to its volume, diversity and

speed. First, the sheer volume of data exceeds traditional storage capacities, particularly in areas such as genomics and astronomy (Makani, 2015). Second, different data formats make management difficult and require flexible frameworks. Finally, rapid data generation requires real-time processing, which is crucial for fields such as meteorology (Dewan & Choudhary, 1992).

Data Quality and Integrity

Maintaining data quality and integrity is crucial in research data management (RDM). The main approaches and challenges are reliable data quality (Peng, 2011) and data integrity (Hasselbring & Kao, 2013).

Data Security and Privacy

Data security and data protection are crucial in research data management (RDM), with the following key issues of security, preserving privacy, legal and ethical considerations, and data governance and compliance: Effective governance frameworks are crucial for monitoring data processing practices. By addressing these challenges, researchers can create a safe and ethical research environment that promotes collaboration and innovation while protecting the rights and interests of data subjects.

Data Sharing and Collaboration

Data sharing and collaboration are critical to scientific progress while respecting copyright and intellectual property rights. The most important aspects and challenges include open science and data sharing, interdisciplinary collaboration, data access and reuse, and data citation and attribution. By promoting data sharing, interdisciplinary collaboration, and facilitating data access and reuse, research data management practices support scientific discoveries through the use of collective knowledge and expertise.

Reproducibility and Transparency

Reproducibility and transparency are critical to maintaining research integrity across disciplines. The most important aspects include reproducibility, transparency, replication, and validation. Prioritizing reproducibility and transparency in research data management strengthens research credibility, promotes collaboration, and accelerates scientific progress.

STRATEGIES FOR EFFECTIVE RESEARCH DATA MANAGEMENT

Research data management is critical to accelerating scientific discoveries through the effective handling of research data (Uhlir, 2012). However, dealing with the increasing amount, complexity, and diversity of data represents a challenge (Corti *et al*., 2014). This article explores strategies to improve reproducibility, transparency, and collaboration in research (Wilkinson *et al*., 2016). Careful planning, including comprehensive data management plans, is critical (National Academies of Sciences, Engineering, and Medicine, 2018). Organizing and documenting data optimizes retrieval and analysis (Mc Coach, Dineen, Chafouleas, & Briesch, 2020). Transparent exchange and collaboration accelerate discovery (Costello *et al*., 2013). Data security, privacy, long-term retention, and accessibility are also crucial (Mayernik *et al*., 2013). Adopting these strategies helps researchers overcome challenges and advance scientific knowledge in various areas.

Data Management Planning

Data management planning is critical to the systematic organization of research data (Wilkinson *et al*., 2016). Researchers should outline data collection methods, storage solutions, sharing protocols, and retention strategies (Corti *et al*., 2014; National Academies of Sciences, Engineering, and Medicine, 2018; Medina *et al*., 2022) This ensures data consistency, security, accessibility, and long-term usability. By incorporating these elements, researchers improve transparency, reproducibility, and collaboration in scientific research (Corti *et al*., 2014). Effective planning helps researchers overcome data challenges and accelerate scientific discoveries.

Metadata Standards and Documentation

Metadata standards and documentation are essential to research data management and provide structure and clarity for researchers. Standardized naming conventions and file structures make data organization and retrieval easier (Kohonen *et al*., 2000). Metadata provides important contextual information about the data and aids in its discovery and interpretation (Corti *et al*., 2014). Thorough documentation of research methods and data processing steps ensures transparency and reproducibility (K. Briney, 2015). The

use of data management tools improves the consistency and accessibility of metadata (Mayernik, 2016). Compliance with metadata standards and comprehensive documentation practices improve data usability and facilitate scientific discoveries (Stodden *et al*., 2016).

Data Storage and Backup

Effective research data management relies heavily on robust data storage and backup systems to ensure the availability, reliability, and long-term retention of data (Haynes, 2004). Researchers accumulate large amounts of data during their projects, which requires careful consideration of storage options such as local systems, network storage, cloud storage, or dedicated research data repositories. Security measures such as encryption and access controls are essential to protect sensitive data from unauthorized access and ensure compliance with privacy regulations and ethical guidelines (Wilkinson *et al*., 2016). Each storage option offers different benefits and considerations regarding accessibility, scalability, security, and cost. Regular data backups are crucial to prevent data loss or corruption. Redundancy across multiple locations reduces risks from hardware failure or human error. Implementing version control mechanisms allows researchers to track data changes over time, ensuring preservation and accessibility of previous versions when necessary (K. Briney, 2015). By strengthening data storage and backup strategies, researchers can minimize the risk of data loss, ensure data availability, and maintain data integrity. These practices increase the reliability and continuity of research projects and allow researchers to focus on scientific analysis and discovery.

Data Organization and Version Control

Efficient data organization and version control are essential components of research data management and facilitate data retrieval, interpretation and collaboration (Shukla, George, Tiwari, & Kureethara, 2022). Given the significant amount of data generated, establishing structured file naming conventions is critical. Standardized, descriptive file names make identification and discovery easier, with the inclusion of metadata providing additional context (Dadam, Lum, & Werner, 1984). Researchers should also develop logical folder structures that reflect the hierarchy of the project and enable efficient data categorization and retrieval while minimizing the risk of data loss (Kowalczyk & Shankar, 2011). Version control is of paramount importance as it allows researchers to transparently monitor data iterations, ensuring the preservation and accessibility of previous versions for reproducibility (Zeng & Qin, 2020). Using dedicated version control systems such as Git or Subversion enables efficient file version management and collaboration between team members, improving data integrity and transparency (K. A. Briney, Coates, & Goben, 2020). Comprehensive documentation of research methods and analysis techniques supports transparency and reproducibility and enables validation and further research (McCormick, Liu, Jomier, Marion, & Ibanez, 2014). Adopting effective data organization and version control practices streamlines data management, promotes collaboration, and increases transparency in scientific research.

Data Sharing and Collaboration Platforms

Data sharing and collaboration are essential components of effective research data management and promote transparency, reproducibility, and scientific progress (Strupler & Wilkinson, 2017). Researchers can use data sharing and collaboration platforms to securely share, access and collaborate on research data. These platforms offer various functions such as data repository services, version control, access controls, and metadata management, thereby covering different research needs and disciplines (Mische *et al*., 2020). One notable platform is the Dataverse project, which facilitates data sharing, publishing, and management with features such as DOI attribution and data citation support (Corti *et al*., 2014). Figshare is another valuable platform that allows researchers to upload and share various research results and provides version control, DOI attribution, and robust metadata management (Powers & Hampton, 2019).Subject-specific platforms such as Dryad and Zenodo target specific research areas and offer seamless integration with other research tools and services(Austin *et al*., 2016). When using these platforms, researchers should consider factors such as data licensing options, privacy settings, and access controls to ensure responsible data sharing and respect privacy and intellectual property rights (d'Aquin, Kirstein, Oliveira, Schimmler, & Urbanek, 2023).

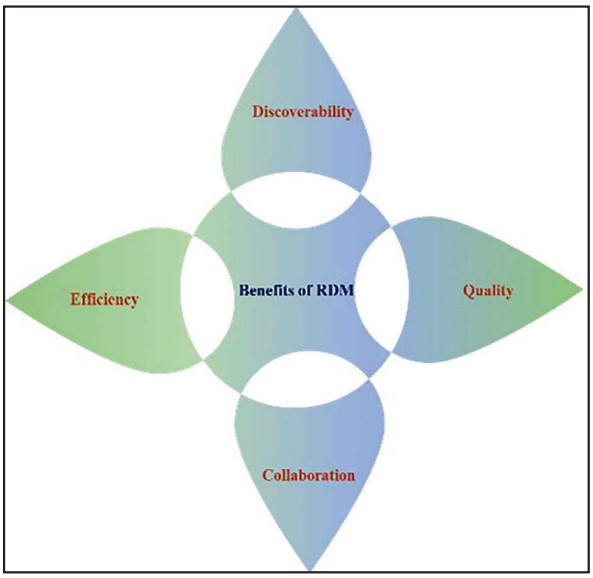
Data Curation and Preservation

Data curation and retention are critical to effective research data management and ensure the ongoing accessibility, usability, and integrity of research data (Mittal *et al*., 2023). Proper curation involves organizing, describing, and preserving data to facilitate to ensure preservation, researchers should select appropriate storage solutions and implement robust backup strategies ranging from local servers to cloud-based systems (Patel, 2016). You should also consider data formats and metadata standards to ensure long-term usability (Kennan & Markauskaite, 2015). Long-term accessibility can be achieved through archiving in specialized repositories that provide persistent identifiers and metadata descriptions (Mazumdar, Seybold, Kritikos, & Verginadis, 2019). Additionally, researchers should address data licensing and rights management to enable appropriate reuse while protecting intellectual property (Doyle, Viktor, & Paquet, 2009). These practices should be integrated throughout the research data lifecycle, starting with planning and collection, to ensure the integrity and longevity of research data (Wilkinson *et al*., 2016).

Data Security and Privacy Measures

Ensuring data security and privacy is crucial for effective research data management and protects the confidentiality, integrity, and availability of data (Amorim, Castro, Rocha da Silva, & Ribeiro, 2017). Researchers must take measures to protect sensitive data and comply with data protection regulations. Access controls are essential and limit access to authorized individuals through authentication mechanisms such as passwords or two factor authentication (Sun, Zhang, Xiong, & Zhu, 2014). Anonymizing data by removing personally identifiable information mitigates risks of re-identification and preserves participant privacy using techniques such as data aggregation or anonymization (Bruzgiene & Jurgilas, 2021). Encryption is critical to securing data during storage and transmission. Compliance with data protection regulations such as the GDPR is essential and requires adherence to legal and ethical standards, including obtaining informed consent and providing opportunities for data subjects to exercise their rights (Gharaibeh *et al*., 2017).

BENEFITS OF EFFECTIVE RESEARCH DATA MANAGEMENT



Effective research data management is critical to accelerating scientific discovery and optimizing the value of research data (Korovessis, 2015). It ensures the accessibility, usability, and value of data and promotes reproducibility, transparency, and collaboration (Corti *et al*., 2014). Managing data comprehensively includes activities such as collecting, storing, organizing, sharing, documenting, and preserving, addressing challenges, and accelerating scientific discovery while enabling validation and replication of findings. Transparent and accessible research data support rigorous analysis and promote open science principles, interdisciplinary collaboration, and societal impact (Borgman, 2015; Wilkinson *et al*., 2016). The benefits of effective data management include improved discoverability, data quality, research efficiency, and collaboration opportunities (Peterka *et al*., 2020). Investments in data infrastructure, policies, and resources contribute to a sustainable research ecosystem (Wilkinson *et al*., 2016).

Enhanced Collaboration and Interdisciplinary Research

Effective research data management significantly promotes collaboration and facilitates interdisciplinary research (Procter, Halfpenny, & Voss, 2012). By implementing robust data management practices,

researchers can collaborate across disciplines and address complex scientific questions by sharing, reusing, and building on well-managed research data (Finkel *et al*., 2020). Collaboration and interdisciplinary research are critical to addressing complex scientific challenges and overcoming obstacles such as data fragmentation and inaccessibility. Well-managed data, properly organized, documented, and shared, improves the visibility and discoverability of research results and makes it easier for researchers across disciplines to identify relevant data sources (Van den Eynden, Corti, Woollard, Bishop, & Horton, 2011). Standardized metadata and documentation help understand data context and properties and facilitate integration into interdisciplinary research projects. Effective research data management ensures the harmonization, consistency, and compatibility of data sets and enables researchers to merge and analyze data across disciplines (Creamer, Martin, Kafel, & Wood, 2014). Thus, effective research data management is critical to improving collaboration and enabling interdisciplinary research. Facilitating the exchange, integration, and analysis of research data promotes collaborative efforts, interdisciplinary approaches, and the exploration of complex scientific questions. These collaborative efforts, underpinned by effective research data management, have the potential to address global challenges, and achieve significant societal impact.

Data Sharing and Reproducibility

Effective research data management is critical to promoting data sharing and reproducibility, which is critical to scientific progress (Kowalczyk & Shankar, 2011). Robust data management practices enable researchers to optimize sharing and collaboration and improve reproducibility. Data sharing enables wider accessibility, transparency, and validation and increases credibility (Feinberg *et al*., 2020). Furthermore, data sharing facilitates the exploration of new research questions by combining data sets, leading to new insights (Sampson *et al*., 2019). Effective data management strengthens reproducibility, a cornerstone of scientific research (Nosek *et al*., 2015). By providing well-documented data, researchers facilitate replication, thereby ensuring robustness (Leonelli, 2013). Data sharing and reproducibility contribute to cumulative scientific knowledge, enabling meta-analyses and encouraging further exploration (Alvarez, Key, & Núñez, 2018). Reproducible findings underpin advances across disciplines (Pasquetto, Borgman, & Wofford, 2019). Effective data management ensures transparency and reliability and facilitates verification, which is crucial for scientific progress. These practices accelerate discovery, encourage collaboration, and expand knowledge across disciplines.

Accelerated Scientific Discovery

Effective research data management is crucial for advancing scientific knowledge through data sharing and reproducibility (Cunha-Oliveira, Ioannidis, & Oliveira, 2024). Robust data management practices enable researchers to optimize sharing and collaboration and improve the reproducibility of results. Sharing data promotes transparency and allows others to validate research findings, leading to more credible results (Miguel *et al*., 2014). Furthermore, data sharing promotes the exploration of new research avenues by combining data sets, thereby facilitating novel discoveries (Keim, Kohlhammer, Ellis, & Mansmann, 2010). Effective data management also strengthens reproducibility, which is crucial for scientific investigations (Nosek *et al*., 2015). Well-documented data enable replication and ensure the validity of research findings (Alston & Rick, 2021). Furthermore, both data sharing and reproducibility contribute to cumulative scientific knowledge, supporting meta-analyses and facilitating further research (Hardwicke *et al*., 2020). Reproducible findings stimulate new hypotheses and advances across disciplines (Tenopir *et al*., 2015). In summary, effective research data management promotes transparency, reliability and collaboration, which are essential for scientific progress.

Increased Research Efficiency and Cost-effectiveness

Effective research data management offers significant benefits in terms of increased research efficiency and cost-effectiveness and leads to more productive scientific investigations (Gatchel & Okifuji, 2006). Robust data management practices streamline workflows and save time, resources, and money. Improved organization and documentation of research data enable quick data access and shorten search time (Gray *et*

al., 2005). Comprehensive documentation improves reproducibility, facilitates validation, and builds on previous work. Seamless data exchange promotes collaboration, promotes interdisciplinary research, and leads to comprehensive results (Wilkinson *et al*., 2016). Reusing data minimizes redundant effort and saves time and resources (Dutch, 2008). Effective data management also enables the use of advanced analysis tools, thereby increasing efficiency. The resulting research efficiency leads to cost savings for researchers and institutions, optimizes resource utilization, and mitigates legal risks (Costello *et al*., 2013).

EMERGING TRENDS AND FUTURE DIRECTIONS

New trends in research data management, including AI integration, cloud computing, improved data sharing, and legal considerations, are shaping the future of scientific discovery (Reyes-Ortiz *et al*., 2018; Borgman, 2015; Kratz & Strasser, 2015; Reilly *et al*., 2018; Dove *et al*., 2019; Mayernik *et al*., 2013). AI and ML automate tasks, cloud computing provides scalable storage and computing resources, while data-sharing policies and collaborative platforms promote interdisciplinary research. Ethical concerns such as privacy and data management are critical, as are strategies for long-term data preservation and sustainability.

Data Integration and Interoperability

Data integration and interoperability, key trends in research data management, are critical to advancing scientific discoveries. Standardized formats and semantic web technologies enable smooth data exchange and interdisciplinary research (Kratz & Strasser, 2015; Auer *et al*., 2018). Integrating data into computational models improves model accuracy and alignment with empirical data (Wilkinson *et al*., 2016). Researchers need robust infrastructure, tools, and collaborative networks to achieve effective data integration (Parsons *et al*., 2019). These trends are driving comprehensive analysis, deeper insights, and scientific advancement.

Artificial Intelligence and Machine Learning

Artificial intelligence (AI) and machine learning (ML) represent promising trends in research data management and are reshaping scientific discovery. They automate tasks, improve data quality, and uncover hidden patterns, supporting data integration and hypothesis generation (Lecun *et al*., 2015; Ching *et al*., 2018; Angermueller *et al*., 2016). AI and ML enable predictive modeling, improve data discovery, and require ethical considerations for responsible use (Lundberg *et al*., 2018; Jha *et al*., 2017; Hutson, 2018). These technologies offer opportunities to accelerate research and gain deeper insights from complex data sets.

Data-Driven Discovery and Predictive Analytics

Data-driven discovery and predictive analytics are promising trends in research data management that are driving scientific research forward. By using advanced computational methods, researchers uncover patterns and trends in large data sets, promoting hypothesis generation and experimentation (Hey *et al*., 2009). Predictive analytics focuses on predicting future outcomes and improving decision-making across scientific fields (Gandomi & Haider, 2015). These approaches streamline research and enable researchers to efficiently generate insights and prioritize avenues for exploration (Ghosh *et al*., 2018). Challenges include data quality, algorithm selection, and ethical considerations that require careful management (Laney, 2015). Overall, data-driven discovery and predictive analytics offer potential to accelerate scientific progress and understand complex research challenges.

Open Science and FAIR Data Principles

Open Science and FAIR Data Principles are transforming research data management and scientific discovery. Open Science values transparency and collaboration, while FAIR principles promote discoverability and reuse of data. Open Science promotes the open exchange of data and methods and improves collaboration and interdisciplinary research (Fecher & Friesike, 2014). FAIR principles ensure careful data description, persistent identifiers and open protocols, facilitating data integration and reuse (Axton *et al*., 2016).

INSTITUTIONAL SUPPORT FOR RESEARCH DATA MANAGEMENT

In the modern scientific landscape, effective management of research data is critical to accelerating discoveries (Jones, Pryor, &

Whyte, 2019). Institutions play a key role in supporting data management (Cox, Pinfield, & Rutter, 2017). This article examines how institutional support affects scientific progress. Research data management includes various activities such as collecting, storing and sharing (Unal, Chowdhury, Kurbanoğlu, Boustany, & Walton, 2019). Institutions provide researchers with important tools and guidelines for effective data management.

Policy Development and Implementation

Institutional support for research data management includes the development of data processing, sharing, and storage policies (Jones, Pryor, & Whyte, 2019). Effective policies are developed collaboratively and cover areas such as data collection, storage, sharing protocols, and ethical considerations (Poline *et al*., 2012). Guidelines require clear communication and training for researchers to ensure compliance. Institutions can provide specialized support services to support researchers (Saupe, 1990). Robust policies promote a culture of data stewardship, ensure data integrity, and contribute to scientific progress.

Training and Education

Institutional support includes training programs to train researchers in effective data management (Flores, Brodeur, Daniels, Nicholls, & Turnator, 2015). These programs cover various aspects and are tailored to the needs and career stages of researchers (Jones, Pryor, & Whyte, 2019). Training improves data integrity and promotes a culture of accountability Institutions can collaborate with external partners to provide diverse expertise Overall, the training contributes to scientific progress by promoting transparent and ethical data management practices.

Data Management Services and Infrastructure

Institutional support for research data management includes the provision of services and infrastructure such as data repositories, cloud storage, and data sharing platforms (Gesing, Antons, Piening, Rese, & Salge, 2015). Repositories adhere to retention and access standards (Jones, Pryor, & Whyte, 2019). Cloud storage offers scalable options and data sharing platforms promote collaboration (Chiware, 2020). Institutions are also investing in analytics and visualization tools that help researchers derive insights from their data (Fabian, Ermakova, & Junghanns, 2015). These resources improve data integrity and accessibility, promoting scientific discovery.

Research Data Management Committees and Support Networks

Institutions often form research data management committees (RDM) and support networks to monitor and support effective data management (Williamson, 2016). These groups, consisting of various stakeholders, develop policies, services and infrastructure for RDM. They ensure compliance with best practices and regulations (Pinfield, Cox, & Smith, 2014). provide access to resources, and offer training programs (Ko & Rossen, 2017). Support networks provide forums for collaboration and knowledge sharing (Perez-Araos, Barber, Eduardo Munive-Hernandez, & Eldridge, 2007). fostering a culture of data management and collaboration.

CONCLUSION

Research data management (RDM) is central to scientific purposes and enables the collection, organization, sharing and storage of data. In today's dynamic research environment, robust RDM practices are critical to driving scientific progress in the face of increasing data volumes and complexities. This article explores the diverse role of RDM in advancing scientific knowledge, addressing challenges, and exploring new trends. Understanding the importance of RDM promotes research integrity, transparency and collaboration, enabling researchers and institutions to efficiently manage the complexities of data management and accelerate scientific discoveries. Research data management will continue to play a central role in shaping scientific discoveries in the future. Emerging trends such as AI integration, cloud computing, and data sharing platforms will increase RDM efficiency. In addition, ethical and legal concerns require responsible handling of data. Recognizing the importance of RDM and investing in resources can promote an open, transparent and collaborative research landscape and accelerate scientific progress.

REFERENCES

1. AlNoamany, Y., & Borghi, J.A. (2018). Towards computational reproducibility: researcher

perspectives on the use and sharing of software. *PeerJ Computer Science*, *4*, e163.

2. Alston, J.M., & Rick, J.A. (2021). A beginner's guide to conducting reproducible research. *Bulletin of the Ecological Society of America*, *102*(2), 1-14.
3. Alvarez, R.M., Key, E.M., & Núñez, L. (2018). Research replication: Practical considerations. *PS: Political Science & Politics*, *51*(2), 422-426.
4. Amorim, R.C., Castro, J.A., Rocha da Silva, J., & Ribeiro, C. (2017). A comparison of research data management platforms: architecture, flexible metadata and interoperability. *Universal access in the information society*, *16*, 851-862.
5. Austin, C.C., Brown, S., Fong, N., Humphrey, C., Leahey, A., & Webster, P. (2016). Research data repositories: review of current features, gap analysis, and recommendations for minimum requirements. *IASSIST quarterly*, *39*(4), 24-24.
6. Borgman, C.L. (2012). The conundrum of sharing research data. *Journal of the American Society for Information Science and Technology*, *63*(6), 1059-1078.
7. Briney, K. (2015). *Data Management for Researchers: Organize, maintain and share your data for research success*. Pelagic Publishing Ltd.
8. Bruzgiene, R., & Jurgilas, K. (2021). Securing remote access to information systems of critical infrastructure using two-factor authentication. *Electronics*, *10*(15), 1819.
9. Chiware, E.R. (2020). Open research data in African academic and research libraries: a literature analysis. *Library Management*, *41*(6/7), 383-399.
10. Coccia, M. (2021). Evolution and structure of research fields driven by crises and environmental threats: the COVID-19 research. *Scientometrics*, *126*(12), 9405-9429.
11. Corti, L., Van den Eynden, V., Bishop, L., & Woollard, M. (Eds.). (2014). Managing and sharing research data: A guide to good practice. Sage.
12. Costa, L.D.F., Cesar, R.M., & Baptista, R.M. (2020). Scientific data management challenges: from volume to value. *Data Science Journal*, *19*(1), 18.
13. Cox, A.M., Pinfield, S., & Rutter, S. (2017). Research data management and libraries: Current activities and future priorities. *Journal of Librarianship and Information Science*, *49*(3), 294-308.
14. Creamer, A.T., Martin, E.R., Kafel, D., & Wood, S.M. (2014). Research data management and the health sciences librarian. *Health Sciences Librarianship*, 252-274.
15. Cunha-Oliveira, T., Ioannidis, J.P., & Oliveira, P.J. (2024). Best Practices for Data Management and Sharing in Experimental Biomedical Research. *Physiological Reviews*.
16. Dadam, P., Lum, V.Y., & Werner, H.D. (1984, August). Integration of time versions into a relational database system. In *VLDB* (Vol. 84, No. 468, pp. 509-522).
17. Dewan, P., & Choudhary, R. (1992). A high-level and flexible framework for implementing multiuser user interfaces. *ACM Transactions on Information Systems (TOIS)*, *10*(4), 345-380.
18. Dove, E.S., Joly, Y., Knoppers, B.M., & Zeps, N. (2019). Knoppers B.M., & Zeps N. (2019). Towards an ethics code for Big Data in health and research. *Nature Reviews Genetics*, *20*(7), 384-388.
19. Doyle, J., Viktor, H., & Paquet, E. (2009). Long-term digital preservation: preserving authenticity and usability of 3-D data. *International journal on digital libraries*, *10*, 33-47.
20. Dutch, M. (2008, June). Understanding data deduplication ratios. In *SNIA Data Management Forum* (Vol. 7).
21. Elgendy, N., & Elragal, A. (2014). Big data analytics: a literature review paper. In *Advances in Data Mining. Applications and Theoretical Aspects: 14th Industrial Conference, ICDM 2014, St. Petersburg, Russia, July 16-20, 2014. Proceedings 14* (pp. 214-227). Springer International Publishing.
22. Fabian, B., Ermakova, T., & Junghanns, P. (2015). Collaborative and secure sharing of healthcare data in multi-clouds. *Information Systems*, *48*, 132-150.
23. Feinberg, M., Sutherland, W., Nelson, S. B., Jarrahi, M. H., & Rajasekar, A. (2020). The new reality of reproducibility: The role of data work in scientific research. *Proceedings of the ACM on Human-Computer Interaction*, *4*(CSCW1), 1-22.
24. Flores, J.R., Brodeur, J.J., Daniels, M.G., Nicholls, N., & Turnator, E. (2015). Libraries and the research data management landscape. *The process of discovery: The CLIR postdoctoral fellowship program and the future of the academy*, *2010*, 82-102.
25. Gatchel, R.J., & Okifuji, A. (2006). Evidence-based scientific data documenting the treatment and cost-effectiveness of comprehensive pain

programs for chronic nonmalignant pain. *The Journal of Pain*, *7*(11), 779-793.

26. Gharaibeh, A., Salahuddin, M.A., Hussini, S.J., Khreishah, A., Khalil, I., Guizani, M., & Al-Fuqaha, A. (2017). Smart cities: A survey on data management, security, and enabling technologies. *IEEE Communications Surveys & Tutorials*, *19*(4), 2456-2501.

27. Gorgolewski, K.J., & Poldrack, R.A. (2016). A practical guide for improving transparency and reproducibility in neuroimaging research. *PLoS biology*, *14*(7), e1002506.

28. Gray, J., Liu, D.T., Nieto-Santisteban, M., Szalay, A., DeWitt, D.J., & Heber, G. (2005). Scientific data management in the coming decade. *Acm Sigmod Record*, *34*(4), 34-41.

29. Hasselbring, W., & Kao, O. (2013). The future of cloud computing in research. *Communications of the ACM*, *56*(11), 68-76.

30. Hey, T., Tansley, S., & Tolle, K. (2009). The fourth paradigm: data-intensive scientific discovery. Microsoft Research.

31. Ioannidis, J.P., Allison, D.B., Ball, C.A., Coulibaly, I., Cui, X., Culhane, A.C., ... & Gopal-Srivastava, R. (2009). Repeatability of published microarray gene expression analyses. *Nature genetics*, *41*(2), 149-155.

32. Jackson, B. (2018). The changing research data landscape and the experiences of ethics review board chairs: implications for library practice and partnerships. *The Journal of Academic Librarianship*, *44*(5), 603-612.

33. Jones, S., Pryor, G., & Whyte, A. (2019). Institutional RDM policy development in the UK: A maturity model. *International Journal of Digital Curation*, *14*(1), 39-58.

34. Keim, D., Kohlhammer, J., Ellis, G., & Mansmann, F. (2010). *Mastering the information age solving problems with visual analytics*. Eurographics Association.

35. Kennan, M.A., & Markauskaite, L. (2015). Research data management practices: A snapshot in time. *International Journal of Digital Curation*, *10*(2), 69-95.

36. Kim, Y., Stanton, J.M., & Golparvar-Fard, M. (2018). Research data management: A review of current practices. *Journal of Business Research*, *88*, 102-113.

37. Kohonen, T., Kaski, S., Lagus, K., Salojarvi, J., Honkela, J., Paatero, V., & Saarela, A. (2000). Self-organization of a massive document collection. *IEEE transactions on neural networks*, *11*(3), 574-585.

38. Korovessis, P. (2015). *Establishing an information security awareness and culture* (Doctoral dissertation, Plymouth University).

39. Kowalczyk, S., & Shankar, K. (2011). Data sharing in the sciences. *Annual review of information science and technology*, *45*(1), 247-294.

40. Leonelli, S. (2013). Integrating data to acquire new knowledge: Three modes of integration in plant science. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, *44*(4), 503-514.

41. Li, L. (2012). *Data quality and data cleaning in database applications* (Doctoral dissertation).

42. Makani, J. (2015). Knowledge management, research data management, and university scholarship: Towards an integrated institutional research data management support-system framework. *Vine*, *45*(3), 344-359.

43. Mauthner, N.S., & Parry, O. (2013). Open access digital data sharing: Principles, policies and practices. *Social Epistemology*, *27*(1), 47-67.

44. Mayernik, M.S. (2016). Research data and metadata curation as institutional issues. *Journal of the Association for Information Science and Technology*, *67*(4), 973-993.

45. Mazumdar, S., Seybold, D., Kritikos, K., & Verginadis, Y. (2019). A survey on data storage and placement methodologies for cloud-big data ecosystem. *Journal of Big Data*, *6*(1), 1-37.

46. McCormick, M., Liu, X., Jomier, J., Marion, C., & Ibanez, L. (2014). ITK: enabling reproducible research and open science. *Frontiers in neuroinformatics*, *8*, 13.

47. Medina, J., Ziaullah, A.W., Park, H., Castelli, I.E., Shaon, A., Bensmail, H., & El-Mellouhi, F. (2022). Accelerating the adoption of research data management strategies. *Matter*, *5*(11), 3614-3642.

48. Meystre, S.M., Lovis, C., Bürkle, T., Tognola, G., Budrionis, A., & Lehmann, C.U. (2017). Clinical data reuse or secondary use: current status and potential future progress. *Yearbook of medical informatics*, *26*(01), 38-52.

49. Miguel, E., Camerer, C., Casey, K., Cohen, J., Esterling, K.M., Gerber, A., ... & Van der Laan, M. (2014). Promoting transparency in social science research. *Science*, *343*(6166), 30-31.

50. Mittal, D., Mease, R., Kuner, T., Flor, H., Kuner, R., & Andoh, J. (2023). Data management strategy for a collaborative research center.

GigaScience, *12*, giad049.

51. Mondschein, C.F., & Monda, C. (2019). The EU's General Data Protection Regulation (GDPR) in a research context. *Fundamentals of clinical data science*, 55-71.

52. National Research Council. (1997). *Bits of power: Issues in global access to scientific data*. National Academies Press.

53. Nelson, G.S. (2015, April). Practical implications of sharing data: a primer on data privacy, anonymization, and de-identification. In *SAS global forum proceedings* (pp. 1-23).

54. Patel, D. (2016). Research data management: a conceptual framework. *Library review*, *65*(4/5), 226-241.

55. Perez-Araos, A., Barber, K.D., Eduardo Munive-Hernandez, J., & Eldridge, S. (2007). Designing a knowledge management tool to support knowledge sharing networks. *Journal of Manufacturing Technology Management*, *18*(2), 153-168.

56. Pinfield, S., Cox, A.M., & Smith, J. (2014). Research data management and libraries: Relationships, activities, drivers and influences. *PLoS one*, *9*(12), e114734.

57. Powers, S.M., & Hampton, S.E. (2019). Open science, reproducibility, and transparency in ecology. *Ecological applications*, *29*(1), e01822.

58. Procter, R., Halfpenny, P., & Voss, A. (2012). Research data management: Opportunities and challenges for HEIs. *Managing research data*, 135-150.

59. Qin, J., Lancaster, F.W., & Allen, B. (1997). Types and levels of collaboration in interdisciplinary research in the sciences. *Journal of the American Society for information Science*, *48*(10), 893-916.

60. Ranganathan, S. (2009). Relevance of research data management and data governance in the context of the three Vs of data-intensive science: Volume, variety, and velocity. In Proceedings of the 9th International Conference on Current Research Information Systems (pp. 371-382).

61. Reyes-Ortiz, J.L., Oneto, L., Samà, A., Parra, X., & Anguita, D. (2018). Transition-aware human activity recognition using smartphones. *Neurocomputing*, *171*, 754-767.

62. Shukla, S., George, J.P., Tiwari, K., & Kureethara, J.V. (2022). Data security. In *Data Ethics and Challenges* (pp. 41-59). Singapore: Springer Singapore.

63. Siedlok, F., & Hibbert, P. (2014). The organization of interdisciplinary research: modes, drivers and barriers. *International Journal of Management Reviews*, *16*(2), 194-210.

64. Singh, R.K., Bharti, S., & Madalli, D.P. (2022). Evaluation of Research Data Management (RDM) services in academic libraries of India: A triangulation approach. *The Journal of Academic Librarianship*, *48*(6), 102586.

65. Stodden, V., McNutt, M., Bailey, D.H., Deelman, E., Gil, Y., Hanson, B., ... & Taufer, M. (2016). Enhancing reproducibility for computational methods. *Science*, *354*(6317), 1240-1241.

66. Strupler, N., & Wilkinson, T.C. (2017). Reproducibility in the field: Transparency, version control and collaboration on the project panormos survey. *Open Archaeology*, *3*(1), 279-304.

67. Sun, Y., Zhang, J., Xiong, Y., & Zhu, G. (2014). Data security and privacy in cloud computing. *International Journal of Distributed Sensor Networks*, *10*(7), 190903.

68. Tenopir, C., Allard, S., Douglass, K., Aydinoglu, A. U., Wu, L., Read, E., & Frame, M. (2011). Data sharing by scientists: Practices and perceptions. *PLoS One*, *6*(6), e21101.

69. Uhlir, P.E. (Ed.). (2012). *For attribution: developing data attribution and citation practices and standards: summary of an international workshop*. National Academies Press.

70. Van den Eynden, V., Corti, L., Woollard, M., Bishop, L., & Horton, L. (2011). Managing and sharing data; a best practice guide for researchers.

71. Wilkinson, M.D., Dumontier, M., Aalbersberg, I.J., Appleton, G., Axton, M., Baak, A., & Mons, B. (2016). The FAIR guiding principles for scientific data management and stewardship. *Scientific data*, *3*(1), 1-9.

72. Williamson, B. (2016). Digital education governance: data visualization, predictive analytics, and 'real-time'policy instruments. *Journal of Education Policy*, *31*(2), 123-141.

73. Wilson, S., Wallis, J.C., & Paton, N.W. (2017). Scientific research data management: Challenges, opportunities, and strategies. *ACM Computing Surveys (CSUR)*, *50*(6), 1-38.

74. Zeng, M.L., Qin, J. (2020). Metadata. United States: American Library Association.