

# Implementing Best Practices in Open Data Sharing and Documentation: A Guide for Research Data Management in Open Science

**Mrs. Dipti J. Fale<sup>1</sup>**

<sup>1</sup>PhD Scholar,  
Sant Gadge Maharaj Arts, Commerce and Science  
College,  
Walgaon, Amravati

**Dr. Sanjay Wagh<sup>2</sup>**

<sup>2</sup>Librarian,  
Sant Gadge Maharaj Arts, Commerce and Science  
College,  
Walgaon, Amravati

## ABSTRACT

Open data sharing and documentation play crucial roles in research data management and open science. Open data sharing and documentation are vital to help researchers increase the efficiency and transparency of their research process. However research data management and open science lack comprehensive and standardized best practices for open data sharing and documentation. The main objective of this study is the adoption of best practices in open data sharing and documentation can elevate research reproducibility, collaboration, and impact within the scientific community. Researchers should prioritize research data management throughout the research process to ensure that their research data is meticulously documented, easily accessible, and reusable for forthcoming research initiatives. This paper suggests the best practice of open data sharing; researchers should comply with data sharing policies and guidelines established by funding agencies and publishers. These policies typically necessitate researchers to make their data openly accessible post-publication to promote transparency and reproducibility. Additionally, by providing training and support for open data sharing and documentation, institutions can help maximize the impact of their data by making it available for reuse in new studies and analyses. Furthermore, best practice is the utilization of standardized metadata for documenting research data. Metadata should encompass details such as data format, variables, and identifiers to facilitate data discovery and reuse.

Keywords: data sharing open science, open data, research data management, and metadata

## Introduction

Open data sharing and documentation are essential components of research data management in open science. By making data openly available to the research community, institutions can promote transparency, reproducibility, and collaboration, which will ultimately boost researchers' investigations. When research data is made openly accessible, it facilitates rapid discovery, enhances transparency, and encourages collaboration among researchers. Nevertheless, successful data sharing details thoughtful consideration of data management practices to guarantee data quality, integrity, and accessibility.

To ensure effective open data sharing and documentation, institutions should adhere to the best practices such as researchers should be familiar with the data sharing policies of institutions and funding agencies, and ensure compliance and alignment with relevant guidelines. Training for open data sharing in research data management and open science should be comprehensive, interactive, and hands-on, providing researchers with the knowledge, skills, and resources they need to effectively share research data and contribute to the open science movement. Institutions should use standardized

metadata to provide essential information about a research dataset, such as its structure, context, and provenance, to enhance the discoverability and usability of their research data.

## 1. Guidelines and practices for open data sharing and documentation

By following guidelines and policies, institutions can create a culture of openness and collaboration, enabling stakeholders to use data for research, decision-making, and public engagement. Following best practices in data sharing and documentation can enhance the credibility, accountability, and impact of open data initiatives.

### 1.1 Key principles and guidelines for open data sharing

Institutions should make all data openly available and accessible to ensure transparency and accountability in research. Data should be easily accessible to researchers, policymakers, and the public without restrictions or barriers. Data should be in formats that are easily interoperable with other datasets, making it easy for researchers to integrate and analyze different data sources. Data should be provided in a form that is easily reusable and can be used for a variety of purposes beyond original research. Data should follow the principles

of being findable, accessible, interoperable, and reusable, as outlined in the FAIR data principles. Appropriate credit and attribution must be given to its creators when data is shared and reused. Open data must be of high quality and accurately represent the research findings to the community. Sensitive or personal data must be handled and shared according to ethical and legal guidelines to protect the privacy of individuals. Institutions should work closely with the broader scientific community and stakeholders to ensure that open data-sharing is consistent with the community's needs and values.

### **1.2 Effectively promote the adoption of open data-sharing practices**

Institutions should provide training and resources to educate researchers about the benefits of open data practices and make data open and accessible. Incentives need to be created by institutions to encourage researchers to adopt open data practices, such as including data sharing requirements in grant applications, providing funding for research data management and sharing activities, or offering rewards to researchers who make their research data openly available. Collaborations with stakeholders should be initiated by institutions to promote open data practices and create a supportive ecosystem for researchers to share their data. This could include working with funding agencies to incorporate data-sharing requirements in grant agreements, partnering with publishers to support data sharing through data repositories, and engaging with research communities to develop community-specific open data-sharing standards. Research data management support services should be established by institutions to support researchers in adopting open data practices, identify areas for improvement, and assess the impact of initiatives. This may include collecting data on researchers' data-sharing activities, conducting surveys to determine researchers' attitudes towards open data practices, and regularly reviewing institutional open data-sharing policies and practices.

### **1.3 Challenges in implementing open data sharing guidelines**

Many institutions may not be aware of the FAIR data principles and their importance in improving research data management and sharing practices. Implementing the FAIR data principles may require investment in research data management tools, infrastructure, and training, which some researchers may not have access to. Making data findable, accessible, interoperable, and reusable can be complex and time-consuming, especially for institutions that are not familiar with research data management best practices. Institutions may be hesitant to make their data openly available due to concerns about data privacy and security,

especially when dealing with sensitive or confidential data. Different research fields may have unique data management needs and requirements, making it challenging to implement the FAIR data principles in a way that is relevant and useful to all institutions. In some research communities, there may be a lack of incentive or support for openly sharing data, which can hinder the implementation of FAIR data principles.

## **2. Training and support for open data sharing**

Provide training programs, workshops, and resources to help institutions improve their research data management and documentation skills. This can include workshops on research data organization, metadata standards, data citation practices, and data-sharing platforms. Institutions can also provide support through research data management services and infrastructure.

### **2.1 The benefits of effective training programs for open data sharing**

The training program should emphasize the benefits of open data sharing, such as increased visibility and impact of research, improved collaboration, and increased reproducibility of research findings. The program should cover best practices for research data management, including data organization, documentation, storage, and sharing. Institutions should learn how to properly format and describe their research data so that it can be easily discoverable and usable by others. Institutions should be introduced to existing data-sharing platforms and repositories where they can deposit their data for long-term access and preservation. Training should include instructions on how to deposit data, assign metadata, and comply with research data-sharing policies. Institutions should be educated on ethical considerations related to open data sharing, including data privacy, confidentiality, and consent. They should also be made aware of legal requirements, such as copyright and licensing issues, that may impact data sharing. Institutions should be trained on how to ensure the security and confidentiality of their data when sharing it openly. This may include encryption, access controls, and anonymization techniques to protect sensitive data. Institutions should learn how to collaborate with others on data-sharing projects including how to establish open data-sharing agreements, assign roles and responsibilities, and maintain communication throughout the open data-sharing process. The training program should include mechanisms for monitoring and evaluating institutions' adherence to open data-sharing practices. This may involve tracking open data-sharing activities, conducting surveys or assessments, and providing feedback to researchers on their progress.

## **2.2 Challenges in training and supporting open data sharing**

Institutions face challenges in adopting open data-sharing practices due to a lack of awareness or understanding of the importance and benefits of sharing data openly. Many institutions may not fully understand the importance of making their data available to the broader research community or may be unsure about the best ways to share open data in a manner that is useful and accessible to others. Additionally, researchers may encounter institutional or disciplinary barriers that prevent or discourage them from openly sharing data. Institutional policies or procedures may make data sharing challenging for institutions or institutions may have concerns about potential negative repercussions, such as plagiarism or competition from other institutions. These issues can be addressed through training that can provide institutions with education and resources on the benefits of open data sharing. Training can also provide practical guidance on how to prepare and share research data effectively. By raising awareness and offering support, training can help institutions overcome perceived barriers and feel more confident about adopting open data-sharing practices. In addition, training can help institutions understand the importance of research data management and documentation. This ensures that their research data is well-organized, accurately annotated, and easily understandable to others. Such practices can increase the utility and impact of the data, making it more likely to be shared and used by other institutions.

## **3. Standardize metadata and data format.**

Standardizing metadata and data formats is crucial to enhance open data sharing in open science and research data management. The development of an interoperable and robust standard facilitates data reuse, improves discoverability, and supports reproducibility in research. This synthesis highlights key aspects of metadata standardization and its implications for open science.

### **3.1 Benefits of standardized metadata and data formats in open data sharing**

Standardized metadata and data formats ensure that different systems and platforms can easily share and exchange data without encountering capability issues. This facilitates seamless integration and collaboration between organizations. Standardized formats help ensure consistency and uniformity in the structure and management of data. This reduces the chances of errors and inconsistencies in data interpretation and analysis. Standardized metadata and data formats make it easier for others to use and reuse data across different projects and applications. This promotes data sharing and can

lead to greater insights and innovation. Standardized formats make it easier for users to search, discover, and access research data from different sources. This improves the overall accessibility of data making it more readily available for analysis and decision-making. Standardized metadata and data formats promote the long-term sustainability of data-sharing initiatives by ensuring that research data remains usable and relevant over time. This helps prevent the obsolescence of research data and the loss of valuable information.

### **3.2 Challenges in adopting standardized metadata practices in open data sharing**

Many institutions may not be aware of the importance of standardized metadata practices or how to implement them effectively in their research processes. Implementing standardized metadata practices can be time-consuming and may require additional resources, such as training and expertise in metadata standards and best practices. Institutions may be reluctant to change their current research data management practices or methodologies, especially if they have been successful in the past. Some research institutions may not prioritize the adoption of standardized metadata practices or provide researchers with the support and resources needed to implement them effectively. Metadata standards can be complex and may vary depending on the field of research, making it difficult for researchers to choose the most appropriate standard for their research data management needs. Institutions may have difficulty ensuring that their metadata practices align with those of other researchers or entities, which can lead to difficulties in exchanging and interoperating data.

By understanding the specific needs and workflows of institutions, metadata standards can be designed to be intuitive and user-friendly. This can help ensure that institutions are more likely to adopt and use standards effectively. By focusing on the functional needs of institutions, metadata standards can be tailored to provide the most relevant and useful information for their research activities. This can help institutions better discover, access, and use research data. By addressing the functional needs of institutions, metadata standards can be designed to facilitate interoperability and open data sharing across different research communities and disciplines. This can help improve collaboration and data reuse. Involving institutions in the development of metadata standards can ensure that the standards are practical and feasible for institutions to implement. This can help increase compliance with the standards and improve metadata quality.

## **Conclusion**

In conclusion, open data sharing and documentation are essential best practices in research data management and open science. By making research data openly available to the public, researchers can maximize the impact of their work, facilitate collaboration and transparency, and promote reproducibility. Good documentation ensures that the data is well-organized, understandable, and accessible to others. Best practices for data sharing and documentation, such as institutions should adopt policies that encourage the sharing of research data. Training on open data sharing is vital for improving data utility for future users. Structured training frameworks for open data sharing are necessary to professionalize their roles and enhance research data management practices in institutions. Standardizing metadata and data formats is important for increasing open data sharing in research data management and open science. The development of interoperable and robust standards facilitates data reuse, improves discoverability, and supports reproducibility in research.

## Reference

- Cunha-Oliveira, T., Ioannidis, J. P., & Oliveira, P. J. (2024). Best Practices for Data Management and Sharing in Experimental Biomedical Research. *Physiological Reviews*. <https://doi.org/10.1152/physrev.00043.2023>
- Open Science Best Practices in Data Science and Artificial Intelligence. (2023). <https://doi.org/10.52825/cordi.v1i.299>
- Ramstrand, N., Fatone, S., Dillon, M. P., & Hafner, B. J. (2020). Sharing research data. *Prosthetics and Orthotics International*, 44(2), 49–51. <https://doi.org/10.1177/0309364620915020>
- Deeb, H., Zieliński, T., & Millar, A. J. (2024). A Protocol for Assessing Open Data Practices: Honours Students Can Lead the Way. <https://doi.org/10.17504/protocols.io.kxygxyxmdl8j/v2>
- Open research data and methods. National policy and executive plan by the higher education and research community for 2021–2025. (2023). *Vastuullisen Tieteen Julkaisusarja*. <https://doi.org/10.23847/tsv.669>
- Michener, W. K. (2013). *Ecological Data Archiving and Sharing*. <https://doi.org/10.1016/B978-0-12-409548-9.10575-5>
- Building Research Data Management (RDM) Expertise and Training Resources in ELIXIR Nodes. (2023). <https://doi.org/10.52825/cordi.v1i.382>
- González-Vélez, H., Dobre, C., Sánchez Solís, B., Antinucci, G., Feenan, D., & Gheorghe, D. (2022). Open Science and Research Data Management: A FAIR European Postgraduate Programme. 2522–2531. <https://doi.org/10.1109/BigData55660.2022.10020931>
- Open Science and Research Data Management: A FAIR European Postgraduate Programme. (2022). 2022 IEEE International Conference on Big Data (Big Data). <https://doi.org/10.1109/bigdata55660.2022.10020931>
- Milčiuvienė, S., Amilevičius, D., & Raškinis, G. (2024). Research Data Management Within the EU Data Policy Framework. *Baltic Journal of Law & Politics*, 17(2), 168–187. <https://doi.org/10.2478/bjlp-2024-00020>
- Langlais, P.-C. (2023). Open Scientific Data. <https://doi.org/10.52949/69>
- Wang, S., Huang, X., Zhang, M., Bao, S., Liu, L., Fu, X., Zhang, T., Song, Y., Kedron, P., Wilson, J. P., Ye, X., Hao, X., & Guan, W. W. (2025). Open science 2.0: revolutionizing spatiotemporal data sharing and collaboration. *Computational Urban Science*, 5(1). <https://doi.org/10.1007/s43762-025-00165-1>
- Cunha-Oliveira, T., Ioannidis, J. P., & Oliveira, P. J. (2024). Best Practices for Data Management and Sharing in Experimental Biomedical Research. *Physiological Reviews*. <https://doi.org/10.1152/physrev.00043.2023>
- Rokem, A., Mandava, V., Cristea, N., Tambay, A., & Connolly, A. J. (2024). Towards an open-source model for data and metadata standards. <https://doi.org/10.31219/osf.io/br6u2>
- Asok, K., Dandpat, S. S., Gupta, D. K., & Shrivastava, P. (2024). Common Metadata Framework for Research Data Repository: Necessity to Support Open Science. *Journal of Library Metadata*. <https://doi.org/10.1080/19386389.2024.2329370>
- Gao, Y. (2024). Development and practices of the national standard for Data paper publication metadata. *China Scientific Data*, 9(1), 1–5. <https://doi.org/10.11922/11-6035.csd.2023.0187.zh>



Alemu, G. (2023). *Metadata Standards and Models*. Elsevier BV. <https://doi.org/10.1016/b978-0-323-95689-5.00035-3>

Kafkas, S., Celebi, R., Ali, M., Jabeen, H., Dumontier, M., & Hoehndorf, R. (2020). *Metadata standards for the FAIR sharing of vector embeddings in Biomedicine*. <https://www.narcis.nl/publication/RecordID/oai%3Acris.maastrichtuniversity.nl%3Apublications%2F6a53b44c-7e5d-430b-bdea-e898be21cea6>

Pradhan, P., & Zala, L. N. (2023). *Role of Research Data Repositories and Open Data Sharing with Specific Reference to Open Science Framework (OSF)*. <https://doi.org/10.17821/srels/2023/v60i6/171159>

Zakharova, O. V. (2023). Scientific documents metadata as a component of the system of the “open science” information resources. *Problemy Programirovaniâ*. <https://doi.org/10.15407/pp2023.04.027>

Gibson, J. (2023). *Sharing, curation and metadata as essential components of the data management plan*. <https://doi.org/10.3233/isu-230220>

Forero, D. A., Curioso, W. H., Patrinos, G. P., & Patrinos, G. P. (2021). The importance of adherence to international standards for depositing open data in public repositories. *BMC Research Notes*, 14(1), 405. <https://doi.org/10.1186/S13104-021-05817-Z>

Silva, F., Amorim, R. C., Castro, J. A., Silva, J. R. da, & Ribeiro, C. (2016). *End-to-End Research Data Management Workflows* (pp. 369–375). Springer, Cham. [https://doi.org/10.1007/978-3-319-49157-8\\_32](https://doi.org/10.1007/978-3-319-49157-8_32)

Borghi, J. A., & Van Gulick, A. (2021). Promoting Open Science Through Research Data Management. *arXiv: Digital Libraries*. <https://dblp.uni-trier.de/db/journals/corr/corr2110.html#abs-2110-00888>

Swedlow, J. R., Kankaanpää, P., Sarkans, U., Goscinski, W., Galloway, G. J., Sullivan, R. P., Brown, C. M., Wood, C. D., Keppler, A., Loos, B., Zullino, S., Longo, D. L., Aime, S., & Onami, S. (2020). A Global View of Standards for Open Image Data Formats and Repositories. *arXiv: Other Quantitative Biology*. <https://arxiv.org/abs/2010.10107>

Hollmann, S., Attwood, T. K., Bongcam-Rudloff, E., Duca, D., D’Elia, D., Endrullat, C., Frohme, M., Messerschmidt, K., & Regierer, B. (2018). *Standardization and Quality Assurance in Life-Science Research - Crucially Needed or Unnecessary and Annoying Regulation?* (pp. 13–20). Springer, Cham. [https://doi.org/10.1007/978-3-030-00825-3\\_2](https://doi.org/10.1007/978-3-030-00825-3_2)

Promoting Open Science Through Research Data Management. (2022). *Harvard Data Science Review*. <https://doi.org/10.1162/99608f92.9497f68e>

Link, G. J. P., Germonprez, M., Goggins, S., Hemsley, J., Rand, B., & Squire, M. (2016). *The Open Community Data Exchange: Advancing Data Sharing and Discovery in Open Online Community Science*. 7. <https://doi.org/10.1145/2962132.2962142>

Latif, A., Limani, F., & Tochtermann, K. (2019). A Generic Research Data Infrastructure for Long Tail Research Data Management. *Data Science Journal*, 18(1), 17. <https://doi.org/10.5334/DSJ-2019-017>