

Perception and Practice of Data Management Plans in Health: An Exploratory Study

Viviane Veiga^{1†}, Luís Ferreira Pires², Patricia Henning¹, João Moreira², Isabella Henrique Lima Pereira^{1,3}

¹Oswaldo Cruz Foundation, Rio de Janeiro-RJ, 21040-900, Brazil

²University of Twente, Enschede 7522 NH, The Netherlands

³Fluminense Federal University, Niterói-RJ 24070-090, Brazil

Keywords: Data management plan; DMP; Researcher perception; DMP tools; DMP template

Citation: Veiga V., Pires L.F., Henning P., et al.: Perception and Practice of Data Management Plans in Health: An Exploratory Study. *Data Intelligence*, Vol. 7, Art. No.: 2025r14, pp. 381–396, 2025. DOI: <https://doi.org/10.3724/2096-7004.di.2024.0030>

ABSTRACT

This paper discusses the quality of Data Management Plans (DMPs) in the health sector and assesses remove the researchers' perceptions of DMPs. We applied qualitative methods to examine publicly available DMPs in healthcare, analyzing researchers' views and practices for creating these plans. The study combines three research methods: analysis of DMPs in the health sector, semi-structured questionnaires, and interviews. Our findings reveal that researchers are generally unaware of the importance and usefulness of DMPs, and acknowledge various inconsistencies and challenges in their development. In this paper, we identified that data management practices need to be improved and advocate for automating them and making DMPs machine-actionable. We also recommend more educational programs, such as workshops and courses, in data management especially for researchers. Finally, we recommend defining clear, accessible guidelines for researchers to effectively elaborate DMPs, and institutionalizing data management within organizations by establishing data (or digital) competence centers.

1. INTRODUCTION

Open Science aims to strengthen scientific rigor, data transparency and reproducibility. However, to achieve these goals, appropriate research data management practices need to be applied. As data-driven research proliferates in academic and scientific environments, advanced tools to support Research Data

[†] Corresponding author: Viviane Veiga (E-mail: viviane.veiga@fiocruz.br; ORCID: 0000-0001-8318-7912).

Management (RDM) must be adopted. In this context, a Data Management Plan (DMP) is an essential tool since it consists of a structured and formal document expected to meticulously outline the roles, responsibilities, and activities in data management. It also addresses the intricacies of data reuse and licensing of research results [1].

DMPs have become mandatory in many universities and research institutions, to meet the demands of financing agencies, confirming the importance of this instrument in the academic and scientific communities. The National Institutes of Health (NIH) has one of the earliest known policies to require a data sharing plan since 2003 [2].

Similarly, the European Commission incorporated DMP requirements under its Horizon 2020 Program [3]. DMPs are expected to comply with the FAIR (Findable, Accessible, Interoperable, Reusable) data principles, encapsulating the philosophy of being “as open as possible, as closed as necessary” [3], while allowing for exceptions under specific circumstances. Since then, DMPs have become an integral part of data management strategies, requiring the definition of templates and tools to assist researchers in their development. As the complexity of research data continues to grow, there is an increasing need for the involvement of different stakeholders in the data management process, as “researchers often lack the necessary knowledge and time to develop and implement a DMP effectively” [4]. This implies that concerns have been raised about the quality and usefulness of the currently available DMPs, which motivated our research. This paper aims at evaluating the quality of DMPs elaborated by researchers in the health domain, especially in Neuroscience. In addition, we assess the perception of these researchers regarding the difficulties encountered during this elaboration, and how they perceive the usefulness of DMPs.

This paper is further structured as follows: Section 2 introduces the theoretical background of this work; Section 3 presents the methodology we applied in our research; Section 4 discusses our analysis of the selected DMPs; Section 5 examines the perception of the researchers who elaborated the DMPs; Section 6 discusses related work; and Section 7 gives presents our conclusions.

2. BACKGROUND

This section covers the background of our research by discussing DMPs and their quality assessment, and introducing the evaluation criteria used in this research.

2.1 Data Management Plans (DMPs)

DMPs play a guiding role throughout all data-related aspects of the stages of research activities, beginning with initial planning and extending beyond the finalization of these activities, making them crucial for effective data management [1].

DMPs have become mandatory in many universities and research institutions to meet the demands of funding agencies as a condition for resource allocation. Coupled with the scientific journals' demands for access to research data, some gaps are highlighted in [5] as: lack of responsibility for effective

implementation of DMPs, deficiencies of data management policies, particularly regarding storage, licensing, metadata and data sharing; and the need for enhanced training and guidance in tool usage.

In this sense, the new generation of DMPs tools incorporate machine-actionable resources that integrate with various scientific systems, facilitating “the continuous and dynamic participation of different stakeholders throughout the management process” [6]. This integration aids researchers in completing templates, streamlining scientific work, and minimizing errors.

The Common Standards DMP Working Group of the Research Data Alliance (RDA) recommends a standard application profile that allows a DMP to be represented as machine-actionable structured data (maDMP) [6].

2.2 DMP quality assessment

By performing exploratory research, we identified few studies in the literature that evaluate the quality and identify challenges associated with the elaboration of DMPs. Notably, the study reported in [7] assessed the quality of DMPs of projects funded by the National Science Foundation. This research revealed that many DMPs provided vague or incomplete descriptions of the management and sharing of research data. These findings confirm the opinion that university libraries should invest more effort in promoting, educating, and supporting the elaboration of DMPs.

In 2022, an evaluation of DMPs of projects funded under the Horizon 2020 Program of the European Commission was carried out [8]. This study indicates that data management is extremely useful, not just as a formal requirement, and concludes that funders should set clear and ambitious Open Science requirements. It suggests having a designated individual to assist researchers with data management, proposing a librarian for this role. A recurring request by the researchers for feedback on the contents of the DMP templates was also identified. Additionally, the study advocates for making DMPs available for long-term preservation and curation.

An analysis of funding organizations identified a critical point: although these organizations require the submission of DMPs, they often provide limited guidance on their specific expectations regarding these plans [9]. Addressing these gaps, some related work suggests enhancing DMP tools and templates to truly aid researchers in the DMP elaboration. In [8], the authors advocate for tools that not only facilitate data management and metadata capture, but also assist the researchers in filling in DMP templates.

A recent study [10] identifies the concern of seeking solutions that can improve the evaluation of DMPs and proposed a toolbox for the automated assessment of machine-actionable DMPs. The authors present methods for automating the evaluation of DMPs, asserting that Machine Actionable Data Management Plans (maDMPs) and semantic technologies increase the potential for automatic evaluation of DMPs.

Therefore, we claim that more systematic studies are necessary to assess the researchers’ adherence to DMPs throughout the entire research lifecycle, as well as to improve the quality and usefulness of

these DMPs. As an initial step, our research seeks to evaluate the quality of DMPs prepared by health researchers, particularly in Neurosciences, and to capture and analyze these researchers' perceptions regarding the difficulties encountered in the preparation of DMPs.

2.3 Evaluation criteria

To assess the quality of DMPs, we used as a reference the "Data Management Plan: Elaboration Guide" [11] developed by the Oswaldo Cruz Foundation in Brazil. This guide is targeted the healthcare sector and is tailored to meet Brazilian legal, ethical, and practice requirements. It was developed based on the predefined topics from the "Practical Guide for International Alignment of Research Data Management" [12] published by the Science Europe initiative, which is considered a reference for DMP development in Europe. Both guides adhere to the FAIR data principles, which defines requirements for improved data management practices.

In this research, we evaluated DMPs based on the following seven topics:

- **Topic 1** (Description of data, collection and/or reuse of data) concerns how data is described, collected or created, processed, and reused.
- **Topic 2** (Metadata, documentation, and data quality) focuses on the metadata standards adopted, whether they will be saved with the data set and the documentation.
- **Topic 3** (Legal and Ethical requirements, codes of conduct) covers the management of ethical issues, codes of conduct, privacy, intellectual property, as well as issues related to access to information and confidentiality.
- **Topic 4** (Storage and backup during the research process) examines how data and metadata backup are performed during and after the research process and how data access and security is managed.
- **Topic 5** (Selection and preservation) looks at which data is selected for retention and sharing, and how long the data is planned to be preserved.
- **Topic 6** (Data sharing and reuse) explores how and when the data is shared and any potential restrictions on its use.
- **Topic 7** (Responsibilities and financial resources) concerns the role that should be responsible for data management, and the resources needed to deliver the plan.

3. RESEARCH METHODOLOGY

The exploratory research reported in this paper has been carried from July 2023 to September 2023. Figure 1 illustrates the research methodology applied in this work, which consists of the 4 steps that we took to achieve our research objectives.

In step 1, we sought for public DMPs in Health and Neuroscience available on the DMPTool^① and DMPOnline platforms^②. These platforms were selected because they are popular for managing research

^① Available at <https://dmptool.org>.

^② Available at <https://dmponline.dcc.ac.uk>.

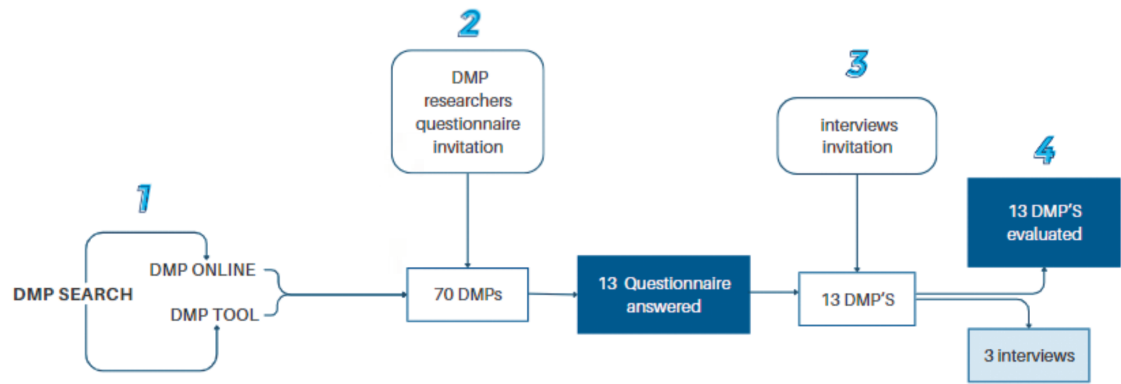


Figure 1. Research methodology.

data and are widely used by the international scientific community. In this step, we identified a total of 70 public DMPs authored by researchers across various health-related disciplines.

In step 2, a semi-structured questionnaire was prepared based on the findings from the literature review, particularly the studies [6, 7, 8, 10] discussed in Section 2. The questionnaire contains 12 questions (see Appendix A) and is aligned with the objectives of this research. The questionnaire was submitted by email to the authors of the 70 public DMPs that were identified in step 1. We received 13 valuable responses.

In step 3, an invitation was sent to the 13 respondents of the questionnaire for an online interview to obtain more accurate information. The interview protocol (see Appendix B) was based on the questions from the questionnaire, allowing open-ended responses and critical discussion. Although only 3 researchers from the 13 accepted our invitation to the interviews, the result of these interviews helped us clarify and confirm some aspects that are discussed in Section 5.

In step 4, we analyzed the level of compliance with our evaluation criteria of the DMPs produced by the 13 researchers. The level of compliance of DMPs was categorized into 3 levels, denoted as + (fully compliant), +/- (partially compliant) and - (non-compliant). In this analysis, we aimed to identify the experience and perception of researchers with the preparation of their DMPs. However, as we identified some inconsistencies in the answers, we sought to find some further clarification during the interviews.

4. DMPS ASSESSMENT

This section discusses the results of the quality assessment of the 13 DMPs in accordance with [11]. Table 1 gives an overview of the DMPs assessment. The 13 researchers are indicated on the left column (rows), while the header (top row) presents the 7 topics that we used for the assessment.

Figure 2 illustrates the coverage (fully, partially, or not compliant) of each topic by the assessed DMPs.

Table 1. Assessment of the DMPs according to [11].

	Topic 1 DESCRIPTION OF DATA, COLLECTION AND/OR REUSE OF DATA	Topic 2 METADATA, DOCUMENTAT ION AND DATA QUALITY	Topic 3 LEGAL AND ETHICAL REQUIREMENTS, CODES OF CONDUCT	Topic 4 STORAGE AND BACKUP DURING THE RESEARCH PROCESS	Topic 5 SELECTION AND PRESERVATION	Topic 6 DATA SHARING AND REUSE	Topic 7 RESPONSIBILITIES AND FINANCIAL RESOURCES
R1	(-)	(-)	(+/-)	(+/-)	(-)	(-)	(-)
R2	(+/-)	(+/-)	(+/-)	(+/-)	(+/-)	(+/-)	(+/-)
R3	(-)	(-)	(+/-)	(+/-)	(-)	(+/-)	(+)
R4	(+/-)	(-)	(-)	(-)	(-)	(-)	(-)
R5	(+/-)	(-)	(+/-)	(+/-)	(+/-)	(+/-)	(+/-)
R6	(+/-)	(-)	(+)	(+/-)	(+/-)	(+/-)	(+/-)
R7	(+/-)	(+/-)	(+)	(+/-)	(+/-)	(+)	(+/-)
R8	(-)	(-)	(+/-)	(+/-)	(+/-)	(+/-)	(-)
R9	(+/-)	(-)	(+/-)	(+/-)	(-)	(+/-)	(-)
R10	(+/-)	(-)	(+/-)	(+/-)	(-)	(-)	(+)
R11	(+/-)	(+/-)	(+/-)	(-)	(-)	(-)	(+/-)
R12	(+/-)	(+/-)	(+)	(+/-)	(-)	(-)	(+)
R13	(+)	(+/-)	(-)	(+/-)	(-)	(+/-)	(+/-)

+ (fully compliant), +/- (partially compliant) and - (not compliant)

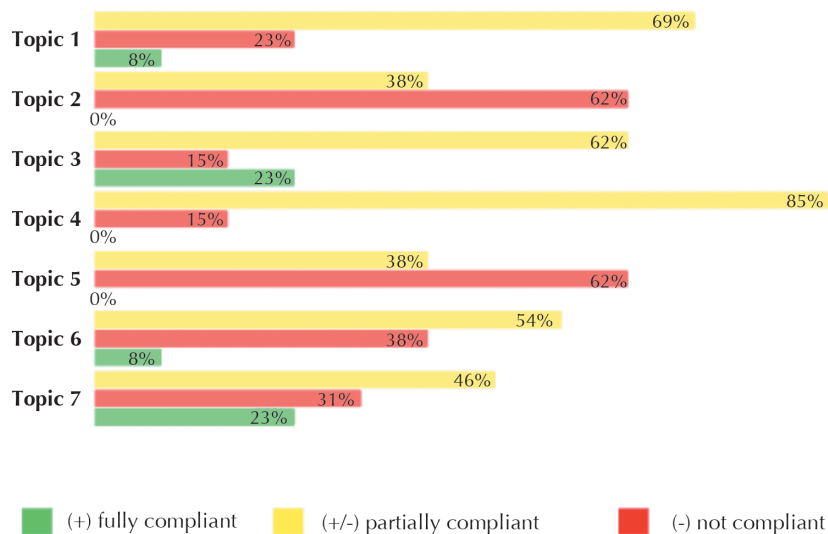


Figure 2. Coverage of the DMPs per topic.

The assessment score that we assigned to each topic is explained below:

- **Topic 1** (Description of data, collection and/or reuse of data). Approximately 69% of the researchers showed partial compliance (+/-) based on the Reference Guide. In contrast, 23% did not provide answers aligned with the Reference Guide, and 8% answered fully correctly. These numbers reveal a significant lack of knowledge regarding the accurate description of data and its reuse.

- **Topic 2** (Metadata, documentation, and data quality). Around 38% of the researchers responded partial compliance (+/-) based on the Reference Guide, 62% did not provide an aligned response, and 23% answered fully correctly. This reveals an unfamiliarity with appropriate metadata description and underscores the importance of investing in guidance in this aspect, crucial for ensuring proper data provenance and reuse.
- **Topic 3** (Legal and Ethical requirements, codes of conduct). Approximately 62% of the researchers responded partial compliance (+/-) based on the Reference Guide, 15% did not provide an aligned response, while 23% answered fully correctly. We observed that legal, ethical, and codes of conduct issues have not been completely clarified. Persistent doubts can lead to significant errors and generate legal complications.
- **Topic 4** (Storage and backup during the research process). About 85% of the researchers answered partial compliance (+/-) based on the Reference Guide, 15% did not provide aligned answers, and there were no fully correct answers. The absence of correct responses regarding storage practices is a concerning factor. This result reveals the urgency of clarifying such practices to avoid potential harm.
- **Topic 5** (Selection and preservation). Around 38% of the researchers responded partial compliance (+/-) based on the Reference Guide, 62% did not provide aligned answers, and there were no fully correct answers. Once again, none of the researchers correctly addressed this topic, and many responses are not aligned with the Reference Guide, so this indicates limited knowledge on the subject.
- **Topic 6** (Data sharing and reuse). Around 54% of the researchers answered partial compliance (+/-) based on the Reference Guide, 38% did not provide aligned responses, and only 2% answered fully correctly. Once again, the accuracy rate was notably low, revealing the pressing need for specific data sharing and reuse guidance.
- **Topic 7** (Responsibilities and financial resources). Approximately 46% of the researchers scored partial compliance (+/-) based on Reference Guide. In contrast, 31% did not provide aligned responses, and 23% answered fully correctly. This topic had the highest rate of correct responses compared to the others. However, the error rate was still quite high, indicating the need for additional guidance to improve understanding and application of these concepts.

Figure 3 shows that only 8.8% of the concerns covered by the topics are accurately addressed. Alarming, 56% of the content in these documents is incomplete, and a significant 35% is either not filled out or filled out incorrectly. This means that a substantial 91% of the information in the DMPs is either incomplete or incorrect, raising serious concerns about their quality, reliability, and effectiveness.

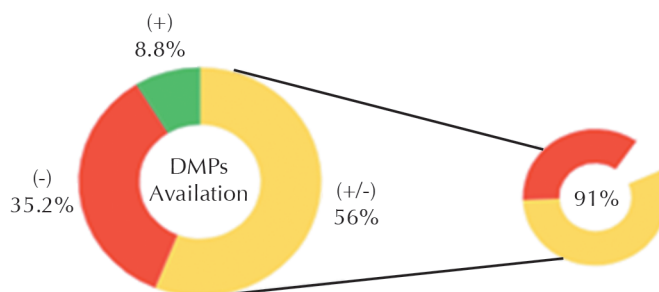


Figure 3. Coverage percentage of the DMPs to the DMPs Guides.

Table 2 provides additional information on the 13 DMPs evaluated, presenting the countries where they were prepared, as well as their respective templates and funding agencies. Two researchers did not mention the institution that funded their research.

Table 2. More information on the 13 evaluated DMPs.

Researcher	Location	Template	Funding Agency
R1	Brazil	State University of Campinas - UNICAMP / Generic Template	Coordination for the Improvement of Higher Education Personnel (CAPES)
R2	Austria	Digital Curation Centre - DCC Template	-
R3	Brazil	Digital Curation Centre - DCC Template	São Paulo Research Foundation (FAPESP)
R4	Brazil	University of São Paulo - USP Template	Coordination for the Improvement of Higher Education Personnel (CAPES)
R5	Scotland	Medical Research Council - MRC Template	Medical Research Council (MRC)
R6	Spain	Digital Curation Centre - DCC Template	-
R7	United States	National Institutes of Health Generic Data Management and Sharing Policy (2023) - NIH-GEN DMSP	National Institutes of Health
R8	United States	National Institutes of Health Generic - NIH-GEN: Generic	Congressionally Directed Medical Research Programs
R9	Brazil	National Institutes of Health Genomic Data Sharing - NIH-GDS: Genomic Data Sharing	National Institutes of Health
R10	Brazil	Digital Curation Centre - DCC Template	São Paulo Research Foundation (FAPESP)
R11	Brazil	Digital Curation Centre - DCC Template	Digital Curation Centre
R12	Brazil	Digital Curation Centre - DCC Template	São Paulo Research Foundation (FAPESP)
R13	United States	National Institutes of Health Generic Data Management and Sharing Policy (2023) - NIH-GEN DMSP	National Institutes of Health

5. PERCEPTION OF THE RESEARCHERS

This section discusses the answers from the 13 researchers to the questions in the questionnaire and the additional insights gained from the interviews. The answers are discussed below referring to the questions in the questionnaire and interview protocol, presented in Appendix A and B, respectively.

(1) Difficulties to elaborate the DMP. Most researchers (11) reported minimal difficulties in creating their DMPs. Only two researchers found the questions confusing or unclear. The main difficulties were the lack of models or integration with national or funding agency specific requirements, and that the questions were found to be too specific and confusing, making it unclear what should be entered in the fields.

(2) Enough instructions to elaborate the DMP. Nine researchers felt that the template instructions were adequate for preparing their DMPs without external help, three researchers found the explanations insufficient, and one had no observations.

(3) Help received to elaborate the DMP. Eleven researchers stated that they did not receive any help in preparing their DMPs, while one received training from their university and another obtained help from an information science specialist to address data protection, archiving, open access publications, and other tasks and requirements related to data management.

(4) DMP's contribution to the organization. Eight researchers found that elaborating a DMP helped organize their data, while five did not have the same perception.

(5) Feedback from funding organizations. Researchers did not receive any form of feedback from the funding organization concerning their DMPs.

(6) DMP template questions. Most researchers (10) found the template questions useful, easy to understand, and the tools simple to use. However, a thorough analysis of the questionnaire responses revealed that several aspects of the template were not fully understood.

(7) DMP tools. Most of the researchers found their DMP tool user-friendly and the DMP template easy to fill in. However, they had several suggestions for improvements, for example, offering more help in the context of each aspect, giving examples that could be reused, possibly from different areas, implementing comment boxes in the tool to take notes, and improving the integration amongst the tools, which are suggestions that may be easier to implement if they are made specific for each domain (e.g., Biology, Social Sciences, Engineering, etc.).

(8) Most difficult questions of the DMP template. Seven researchers found the questions regarding metadata the most challenging. One of the interviewees explained that this difficulty arises from a lack of practice, as they are trained to think in terms of research context, objectives, methods, and anticipated outcomes, but not data management. Four researchers encountered difficulties with issues related to ethics and legal compliance and preservation. A suggestion was made to have the help of ethics and legal experts to fill in the template in a meaningful way. Regarding the preservation topic, one interviewee

mentioned that it was one of the most challenging questions since they had to be creative and devise preservation options, adding that laws are not always clear to scientists.

(9) DMP final version storage and access. Five researchers stated that they will leave the final version of the DMP in a public access repository, four researchers planned to place it in a data repository together with the data set, one researcher considered both options, one researcher did not intend to make their DMPs public, while another one specified a local network location to store their DMP. One of the interviewees suggested that institutions should provide a safe data repository for data storage and backup, avoiding data loss in case a (personal) computer is lost or crashes.

(10) Perception on data management after elaborating the DMP. Nine researchers viewed DMP preparation as an obligation merely to meet funding requirements, while four researchers believed in the practical data management benefits of having a DMP.

(11) Improvement suggestions for DMP templates and tools. Suggestions for improving tools and templates included providing detailed instructions to fill in the templates for different domains and data types, assigning persistent identifiers to data, software and algorithms, and implementing a validation mechanism for repository submissions. One participant found the DMP tool adequate, but the requested content problematic.

(12) Data management and reuse. Almost all researchers (12) agreed that DMP preparation helps in data management and sharing, enhancing the potential for future research reuse. One interviewee mentioned that participating in this study motivated them to seek data management training.

6. RELATED WORK

This section briefly discusses some related work that reflects the growing concerns of stakeholders and researchers improving the quality of DMP tools and templates. Some methods for automating the evaluation of DMPs are explored in [10], highlighting the requirements demanded by DMP reviewers, such as DMP integrity, viability, the quality of proposed actions, and the elimination of ambiguity. This paper also analyzes different scenarios for the automated assessment of DMPs, extending beyond mere review by funders and encompassing contexts where assessment serves as continuous feedback for DMP creators. The paper concludes that specific reviewer tasks can be automated according to the mechanized action capability of DMPs, using semantic web technologies, semantic data integration and knowledge graphs, as well as tools for assessing compliance with FAIR principles. Different from [10], our research covers the analysis of researchers' perception regarding the DMP elaboration.

Fourteen open DMP tools were evaluated in [13], aiming to assist researchers and data managers in creating more effective DMPs through a detailed assessment based on 45 parameters. The study highlights a lack of awareness and implementation of DMPs, especially in Asian and African countries compared to North America and Europe. These tools also fall short in providing adequate support for the FAIR principles, emphasizing the need for improvements in this area.

Similar to ours, this study targets the quality of DMPs but from the perspective of the specific DMP tools and not the users' perception of using them.

A series of case studies from various institutions worldwide is presented in [5], shedding light on current practices and sharing recommendations for future work. The paper concludes with practical recommendations and guidelines to enhance the utility of DMPs, and to facilitate more effective data sharing. Like our research, this study demonstrates concerns with data sharing that can contribute to the improvement of DMP usage. The degree of compliance of ten European DMP templates with the FAIR principles is critically assessed in [14]. The study concludes that among the ten templates analyzed, the "Practical Guide to the International Alignment of Research Data Management" [12] developed by Science Europe achieved the highest degree of fairness. This article was important for our research as we used, in our evaluation, the "Data Management Plan: Elaboration Guide" which was based on the "Practical Guide to the International Alignment of Research Data Management", offering valuable recommendations for the field.

7. CONCLUSIONS

Our study identified some critical and sometimes conflicting insights concerning DMPs and how they are elaborated and perceived. While most researchers reported minimal difficulty in creating their DMPs, their opinions regarding the clarity and specificity of the questions in DMP templates varied. Additionally, a frequent lack of support for DMP preparation was indicated. For instance, one researcher inconsistently stated that "The data will be preserved until the thesis defense date," while also claiming that "The data will be stored indefinitely". This aligns with the observation of some authors about the prevalent lack of knowledge among researchers for adequately preparing and implementing DMPs, highlighting the need for clearer guidance for DMP elaboration [5].

Our study also identified challenging aspects of DMPs, including metadata description, ethical and legal considerations, and issues related to data sharing. In research involving human beings, e.g., in life and social science, often only the ethics committee approval is mentioned, neglecting explicit consent for data sharing or reuse.

Confusion was also noted regarding "intellectual property rights of data", which sometimes mistakenly equated with "rights over scientific publications". Overall, the DMP is acknowledged as a valuable tool for improving data management, sharing and reuse. However, there is an urgent need for researchers to understand the significance of these practices.

Consequently, we highlight the importance of establishing data (or digital) competence centers in an institutionalized way to support Research Data Management, providing data stewardship and information science services, including detailed guidance for DMP elaboration. In case this role is institutionalized, researchers should spend much less time elaborating DMPs. We agree with [15] that globally, research institutions are establishing data competence centers to offer assistance, although this initiative is still in its early stages.

Initiatives aimed at enhancing DMPs should focus on empowering data stewards and researchers through data competence centers. These efforts should equip them with the necessary instruments and knowledge to navigate the complexity and demands of data management within research projects. In turn, this will lead to a more efficient, collaborative, and transparent research environment, yielding benefits for both the scientific community and society.

AUTHOR CONTRIBUTIONS

All the authors contributed collaboratively. Viviane Veiga (viviane.veiga@icict.fiocruz.br), Patricia Henning (henningpatricia@gmail.com) and Isabella Pereira (isabellalima@id.uff.br) were responsible for the design and implementation of the questionnaire, collection, and interpretation of the data. Luís Ferreira Pires (l.ferreirapires@utwente.nl), Viviane Veiga (viviane.veiga@icict.fiocruz.br) and João Moreira (j.luizrebelomoreira@utwente.nl) played roles in structuring the article. Isabella Pereira (isabellalima@id.uff.br), Patricia Henning (henningpatricia@gmail.com), Luís Ferreira Pires (l.ferreirapires@utwente.nl) and Viviane Veiga (viviane.veiga@icict.fiocruz.br) dedicated themselves to the creation of tables and images. Luís Ferreira Pires (l.ferreirapires@utwente.nl), Patricia Henning (henningpatricia@gmail.com), Viviane Veiga (viviane.veiga@icict.fiocruz.br), and João Moreira (j.luizrebelomoreira@utwente.nl) were involved in the description, analysis, and review of the text. Isabella Pereira (isabellalima@id.uff.br) carried out the review of citations and references, in accordance with the journal's standards.

REFERENCES

- [1] Michener, W. K.: Ten simple rules for creating a good data management plan. *PLoS Computational Biology* 11(10), 1–9 (2015). <https://doi.org/10.1371/journal.pcbi.1004525>
- [2] National Institutes of Health: Final NIH statement on sharing research data. <https://grants.nih.gov/grants/guide/notice-files/NOT-OD-03-032.html>. (2003). Accessed 15 January 2023
- [3] European Commission Directorate-General for Research & Innovation: H2020 Programme: guidelines on FAIR data management in Horizon 2020. Version 3.0. Available at: https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf (2016)
- [4] Miksa, T., Cardoso, J., Borbinha, J.: Framing the scope of the common data model for machine-actionable Data Management Plans. In: 2018 IEEE International Conference on Big Data (Big Data), pp. 2733–2742 (2018)
- [5] Bishop, B. W., et al.: Data management plan implementation, assessments, and evaluations: implications and recommendations. *Data Science Journal* 22(27), 1–8 (2023). <https://doi.org/10.5334/dsj-2023-027>
- [6] Miksa, T., Walk, P., Neish, P.: RDA DMP common standard for machine-actionable data management plans. Zenodo (2020). <https://doi.org/10.15497/rda00039>
- [7] Van Loon, J. E., Akers, K. G., Hudson, C., Sarkozy, A.: Quality evaluation of data management plans at a research university. *IFLA Journal* 43(1), 98–104 (2017) <https://doi.org/10.1177/0340035216682041>
- [8] Spichtinger, D., Siren, J.: The development of research data management policies in Horizon 2020. In: *Research Data Management-A European Perspective*, pp. 11–23 (2017)

- [9] Dietrich, D., Adamus, T., Miner, A., Steinhart, G.: De-Mystifying the data management requirements of research funders. *Issues in Science and Technology Librarianship* (70), (2012). <https://doi.org/10.29173/istl1556>
- [10] Miksa, T., et al.: Towards a toolbox for automated assessment of machine-actionable data management plans. *Data Science Journal* 22(28), 1–13 (2023). <https://doi.org/10.5334/dsj-2023-028>
- [11] Veiga, V. S. O., et al.: Plano de gestão de dados de pesquisa - PGD: guia de elaboração. Fiocruz/Icict, Rio de Janeiro (2022) Available at: https://www.arca.fiocruz.br/bitstream/handle/icict/54805/Guia%20PGD%20245x195%20ingles_OK.pdf?sequence=5&isAllowed=y. Accessed 15 January 2023
- [12] Science Europe: Practical guide to the international alignment of research data management-extended edition. Zenodo (2021) <https://doi.org/10.5281/zenodo.4915862>
- [13] Gajbe, S. B., Tiwari, A., Gopalji, Ranjeet K. S.: Evaluation and analysis of data management plan tools: a parametric approach. *Information Processing & Management* 58(3), 1–17 (2021). <https://doi.org/10.1016/j.ipm.2020.102480>
- [14] Henning, P., et al.: The FAIRness of data management plans: an assessment of some European DMPs. *Revista Eletrônica De Comunicação, Informação & Inovação Em Saúde* 15(3), 722–735 (2021). <https://doi.org/10.29397/reciis.v15i3.2270>
- [15] Mons, B.: Invest 5% of research funds in ensuring data are reusable. *Nature*, London 578, 491–491 (2020). <https://doi.org/10.1038/d41586-020-00505-7>

APPENDIX A. QUESTIONNAIRE AND ANSWERS FROM THE RESEARCHERS

1	Did you have any difficulties in preparing your DMP? <i>(Closed)</i> No (11) Yes (2)
1.1	Please, explain the difficulty <i>(Open-ended)</i> • Even though DMPtool was suggested by USP repository as a guide for writing a DMP, it lacks models or integration with Brazilian and/or funding agency specific requirements. For NIH requirements, several models are pre-filled. • Questions were found specific and confusing. They weren't sure what to put in the fields.
2	The DMP instructions were enough to help you prepare you DMP? <i>(Closed)</i> No (3) Yes (9) No difference (1)
3	Did you receive any help from someone to prepare the DMP? <i>(Closed-ended)</i> No (11) Yes (2)
3.1	Who helped you and what kind of help? <i>(Open-ended)</i> • I have received some training from my university. • My research unit has an information scientist specialist. They deal with data protection, archiving, open access publications, and any other tasks and requirements related to data management.
4	Do you think preparing the DMP contributed to your data organization? <i>(Closed)</i> No (5) Yes (8)
5	Did you receive any feedback from the institution that funded the research project related to the DMP? <i>(Closed)</i> No (13) Yes (0)
6	What did you think of the template questions? <i>(Closed)</i> They were useful in dividing the plan into sections but there was not a great explanation of the purpose of each section. (1) Useful and easy to understand. (10)
7	What did you think of the tool used to prepare your DMP? <i>(Closed)</i> User-friendly and easy to fill in. (13) Unfriendly and difficult to fill in. (0)
7.1	Do you have any suggestions for improvement? <i>(Open-ended)</i> • More in-context help would be useful. • I really like the examples and the fact that there are comment boxes on the side of DMPTool, so that I can take notes for myself or collaborators. For the examples, it would be neat if you could select examples in different fields, i.e. immunology, epidemiology, etc. It would be helpful for non-clinical/genomic folks. • Should be more integrated. Can be better done if made specific to domain (Bio, social sciences, engineering, human participants, genomic, clinical)

8 Which parts of the Template were the most difficult to answer? (Tick the required options). (Closed)

Questions about data collection and/or creation. (3)
Questions regarding to metadata. (7)
Questions about Ethics and Legal Compliance. (4)
Questions about storage and backup. (1)
Questions regarding preservation. (4)
Questions about sharing. (3)
Questions about responsibility and resources. (1)
Describing project aims and context. (0)

8.1 Report the reason for the difficulties (Open-ended)

- The time of writing a research grant proposal it might be difficult to estimate aspects like storage and resources needed. Additionally, ethics and legal experts are needed to fill in the template in a meaningful way.
 - I'm not well practiced in selecting the metadata that I'll need up front.
 - Laws are not always clear to scientists.
 - Each country and organization may have a different specific set of requirements. It is very hard to navigate across and DMPtool could be used by organizations to set our specific requirements.
 - Due to lack of habit of doing this.
 - I was not used to talking about my research in these ways.
-

9 Where did you store or intend to store the final version of your DMP at the end of the research? (Closed)

In a repository, along with the datasets. (4)
Leave in the DMPTool or DMPOnline in public access. (5)
I will not make the final version available. (1)
Other: Both in the repository and DMPTool. (1)
Other: My research unit has specific internal network location where we store DMPs. (1)

10 What is your perception on data management after finishing the preparation of the DMP? (Closed)

I prepared the DMP just to satisfy the funding agency. (9)
Helped me manage my data. (4)
A waste of time. (0)

11 Do you have any suggestions to improve the DMP template and tool? (Open-ended)*

- Clearer questions and possible filling models.
 - I like the sample questions/answers.
- Can be made more specific to problem domains and datatypes. It is too general now, which actually makes it harder to fill in, and less useful. Maybe we can mandate that all data and software/algorithms needed for replicating the research must be assigned persistent identifiers and stored in a specific place. This is already suggested somewhat, but we can make it more precise. "There also has to be some enforcement mechanism, maybe a validation schema that makes sure that data and products submitted to repositories satisfy that schema, and will be rejected if not in compliance," This will make it all research data vastly more searchable and useful.
- The platform is fine, the difficulties were in the content.
-

12 In your opinion, the DMP preparation can support the management and sharing of your data and enhance its reuse in new research? (Closed-ended)

Yes (12)
No (1)

APPENDIX B. INTERVIEW PROTOCOL

1	Did you have any difficulties in preparing your DMP?
2	Were the DMP tool instructions enough to help you prepare your DMP?
3	Did you receive any help from someone to prepare the DMP?
4	Who helped you and what kind of help?
5	Do you think preparing the DMP contributed to your data organization?
6	Did you receive any feedback from the institution that funded the research project related to the DMP?
7	What did you think of the tool used to prepare your DMP?
8	Do you have any suggestions for improvement?
9	Which parts of the Template were the most difficult to answer?
10	Where did you store or intend to store the final version of your DMP at the end of the research?
11	What is your perception on data management after finishing the preparation of the DMP?
12	Do you have any suggestions to improve the DMP template and tool?
13	In your opinion, can the DMP preparation support the management and sharing of your data and enhance its reuse in new research?