

# Workshop Report: Developing the NFDI DMP Template Framework

DMP4NFDI is the basic service in development for data management planning within NFDI. The service supports NFDI consortia providing services for data management plans (DMPs), software management plans (SMPs), or analogues. It offers RDMO hosting, support & training for consortia staff and develops and maintains the NFDI DMP Template Framework to facilitate template development and interoperability. The service also supports consortia to integrate DMPs, SMPs, or analogues with their service portfolio.

Within the workshop "Creating Interoperable RDMO Catalogs with the NFDI DMP Template Framework" held on March 12th, 2025 and offered in the context of the E-Science Tage conference in Heidelberg, participants were introduced to the work of the DMP4NFDI service and the first version of the framework. We chose this event to engage with the broader RDM community, since the initial draft was developed in exchange with NFDI consortia and NFDI working groups. The aim was to facilitate discussion, explore its integration into RDMO and to fine tune its content.

In the first part of the workshop, the initial version of the framework and its development was presented. Subsequently, participants engaged in discussions about specific modules of the framework draft. The groups critically discussed these modules according to the three key aspects: content, structure, and transferability. The goal was to identify if all essential aspects are covered adequately and clearly, and if the framework is flexible enough to cover existing DMP templates or use cases. The groups identified several key points and areas for enhancement within the respective modules, which are outlined in the following sections.

## The NFDI DMP Template Framework

The NFDI DMP Template Framework (TF) supports consortia in developing interoperable DMP templates by providing a common vocabulary, a defined structure and a set of questions for DMPs, as well as standardised value ranges for elements necessary to establish interoperability. Its aim is to promote the implementation of machine-actionable DMPs (maDMPs) across NFDI as discussed by the cross-consortia DMP working group *infra-dmp*.<sup>1</sup> DMP4NFDI started its work on this standardisation effort in June 2024. The framework builds upon the international de-facto standard for machine-actionable DMPs<sup>2</sup> developed by a working group of the international Research Data Association (RDA), and has been further adapted to align with commonly used DMP templates in Germany. Workshops within NFDI and exchanges with consortia have shaped the first draft that has now been discussed with the broader RDM community. In the following, we present requirements, suggestions, and ideas discussed for each module of the framework.

---

<sup>1</sup> Diederichs et al. (2024).

<sup>2</sup> Miksa, Walk, & Neish (2020)

## Module: Project

This module covers all project-specific questions, whereby the term “project” could also be replaced with campaign, event, research group, etc. It forms the context in which the DMP is completed and contains questions on overarching, project-related topics, persons, and responsibilities.

- *Management of project information:* It should be possible to store e.g. a general alternative contact person, in case the primary contact becomes unavailable. Additionally, the framework should record the date when contact details were last updated, ensuring transparency and traceability in the event of any changes to the project.
- *Traceability and referenceability of project:* PID providers (e.g., ROR, ORCID) should be mapped or implemented for funder and grant IDs.
- *Clarifications on granularity:* It should be clarified at which points a distinction between projects and subprojects is meaningful and necessary.

## Module: DMP

This module requires generic information about the DMP itself, but also specific information on handling data management tasks. In particular, it defines responsibilities for data management activities. This applies to both the handling of research data and its documentation, as well as more advanced tasks such as metadata management, creating backups, etc.

- *Traceability and referenceability of DMPs:* Persistent Identifiers (PIDs) should be integrated into the framework to ensure clear identification and traceability of DMPs.
- *Documentation and integration challenges:* The metadata schema asked in DMPs must be documented clearly, and the framework should facilitate the collaborative creation, updating, and retrieval of such documentation.

## Module: Dataset

Dataset questions differ from generic questions in that users first had to divide their data into smaller units and answer the corresponding questions for each unit.

This module covers all aspects related to the datasets that are collected or utilised during the research outlined in the DMP. There is no thematic classification here, but rather we look at which sub-areas of data management need to be answered in such detail that they have to be answered at the dataset level.

- *Representation of data processing workflow:* It should be possible to indicate for each dataset how it is related to another dataset in the project, e.g. for pre- and post-processing steps or raw and processed data.
- *Handling of sensitive data:* Sensitive data is more than personal data. Ensure that conditional questions can be answered also for other sorts of sensitive data.
- *Date for data availability:* Entering a specific date for dataset availability is in many cases, especially during the planning phase, not possible. If a date format is needed

for documentation, maybe a two-step solution might work, e.g. a free text question first, if data will be published at all and when (during or at the end of the project, after data collection). And for a later project phase, a question to add the actual date of publication.

- *Data types*: There should be two questions for general information on the dataset, e.g. one question on the general format (texts, images etc.) based on a standardised vocabulary, and a second question on the data format. Both questions with option sets.
- *Clarifications on data publication and granularity*: It has to be clarified whether data publication is intended at project or dataset level. It should also capture the expected data volume, including estimated annual growth.

## Modules: Distribution/Host

The distribution module covers aspects related to data access and publication. In the RDA maDMP standard, several distributions relate to one dataset, but each distribution has just one host. In contrast, in the TF it is also possible to have many hosts within each distribution. The host is the storage location of the distribution and can therefore be splitted into multiple locations depending on the distribution.

- The question *Where will the data (including metadata, documentation and, if applicable, relevant code) be stored or archived after the end of the project?* should be shifted into the host chapter.
- If there are multiple hosts, they must also be described with multiple metadata. There may also be different licensing of the data/metadata.
- The topic of long-term archiving has so far only been addressed superficially. Current guidelines should be made available, and the issue of archiving should be addressed across all data — both published and unpublished data.
- *Access vs. Terms of Use*: Some questions are interdependent. The TF should ensure that dependent questions are only asked if the preceding questions have been answered accordingly, e.g. publication-related questions like "Do you anticipate any implications or restrictions regarding subsequent publication or accessibility?" should not be shown, if data access has previously been declined.

## General Feedback

Some feedback collected concerns the whole framework and its application:

- Different project types (e.g., Collaborative Research Centres) and their specific metadata requirements should be effectively represented in the TF.
- Information regarding expected data volumes, including annual growth rates, should be systematically captured at one central place within the TF.
- The TF should use common RDMO attributes to ensure that existing catalogues can be mapped.
- The framework must remain flexible enough to integrate existing templates. Additionally, best practices for mapping project- or domain-specific DMPs should be developed and included.

# Overall Conclusions and Next Steps

All discussions provided valuable insights for the further development of the NFDI DMP Template Framework. Issues related to flexibility, expandability, and specific metadata requirements were particularly raised. Future steps should focus on e.g. developing concrete solutions for integrating PIDs and terminologies, improving documentation, and facilitating the integration of existing catalogs. A separate exchange will be held with the Long-term Archival (LTA) working group<sup>3</sup> of the NFDI Common Infrastructure Section for a detailed section on digital preservation. We thank all participants for their valuable contributions and feedback, and the fruitful discussions during the workshop.

## Acknowledgements

The DMP4NFDI service is developed as a service in Base4NFDI. Base4NFDI is funded by DFG, as part of [NFDI](#), under project number [521466146](#).



## References

Diederichs, K., Krause, C., Lemaire, M., Reidelbach, M., & Windeck, J. (2024). A Vision for Data Management Plans in the NFDI. <https://doi.org/10.5281/zenodo.10570654>

Markus, K., Leinen, P., & Stäcker, T. (2025). Concept for Setting up an LTA Working Group in the NFDI Section "Common Infrastructures". <https://doi.org/10.5281/zenodo.14939132>

Miksa, T., Walk, P., & Neish, P. (2020). RDA DMP Common Standard for Machine-actionable Data Management Plans. <https://doi.org/10.15497/rda00039>

---

<sup>3</sup> Markus, Leinen, & Stäcker (2025)