A Vision for Data Management Plans in the NFDI

Authors

Katja Diederichs, Celia Krause, Marina Lemaire, Marco Reidelbach, Jürgen Windeck

Version 1.0

January 2024

Table of contents

Abstract

At present, data management plans (DMPs) are still often perceived as mere documents for funding agencies providing clarity on how research data will be handled during a funded project, but are not usually actively involved in the processes. However, they contain a great deal of information that can be shared automatically to facilitate active research data management (RDM) by providing metadata to research infrastructures and supporting communication between all involved stakeholders. This position paper brings together a number of ideas developed and collected during interdisciplinary workshops of the Data Management Planning Working Group (infra-dmp), which is part of the section Common Infrastructures of the National Research Data Infrastructure (NFDI) in Germany. We present our vision of a possible future role of DMPs, templates, and tools in the upcoming NFDI service architecture.

Derzeit werden Datenmanagementpläne (DMPs) häufig noch als reine Dokumente für Förderorganisationen wahrgenommen, die Klarheit über den Umgang mit Forschungsdaten während eines geförderten Projekts schaffen, aber in der Regel nicht aktiv in die Prozesse eingebunden werden. Sie enthalten jedoch eine Vielzahl von Informationen, die automatisch weitergegeben werden können, um ein aktives Forschungsdatenmanagement (FDM) zu ermöglichen, indem sie Metadaten für Forschungsinfrastrukturen und -systeme bereitstellen und die Kommunikation zwischen allen beteiligten Akteuren unterstützen. Dieses Positionspapier fasst einige Ideen zusammen, die während interdisziplinärer Workshops der Arbeitsgruppe Data management planning (infra-dmp), die Teil des Bereichs Common Infrastructures der Nationalen Forschungsdateninfrastruktur (NFDI) in Deutschland ist, entwickelt und gesammelt wurden. Wir präsentieren unsere Vision einer möglichen zukünftigen Rolle von Datenmanagementplänen, -vorlagen und -werkzeugen in der zukünftigen NFDI-Servicearchitektur.

Motivation and Methods

This position paper is the result of workshops and regular meetings of the Data Management Planning Working Group (infra-dmp) within the NFDI section Common Infrastructures during 2023. DMPs are considered to be an important cornerstone of the research data infrastructure, especially in the required cooperation of all partners involved in RDM. They are therefore at the center of this collaboration. The overall goal of our group is the consolidation of existing DMP approaches across all consortia to establish a common understanding and standards for DMPs in the NFDI.

Meanwhile, a large number of useful templates have been developed not only in German-speaking countries but also internationally and at the European level. Initially, research funding bodies were particularly interested in the submission of DMPs for research projects for reasons of proposal review, but it's important to note that the information structurally stored in a DMP has the potential to be used to embed the DMP in a practicable operational concept that combines project and data management as well as documentation and report functions. For this reason, it is important to move away from viewing DMPs solely from the perspective of a supplement to grant proposals and to consider other stakeholder groups and communities with their specific workflows beyond pure research processes. Some concepts, e.g. the machine-actionable DMP (maDMP), have been around for a number of years but their implementation has been slow.[1] The NFDI and the collaboration of the participating consortia now offer the opportunity to promote this more data-centric view and to define interfaces for DMPs in the NFDI service architecture. It is equally important to work together, not only on a purely professional level within an NFDI consortium, but also across consortia, to gather and identify feedback and needs as comprehensively as possible, to identify overlaps and similarities between different communities, and as a result to design generally applicable DMP products and services that can be adapted incrementally to specific user interests.

As the potential uses of DMPs expand, so do the requirements for the content, formal and technical properties of DMP templates. The consortia will establish a common ground for templates and modular extensions, taking into account the interaction of tools and services. With two workshops in 2023, the infra-dmp working group discussed DMP templates and best practices in the consortia, and collected ideas for the future role of DMPs in the NFDI. During the second workshop, the following key questions were explored:

- How can DMPs be made more effective and efficient so that the benefits and motivation to use a DMP continue to increase?
- What opportunities do DMPs offer to move smoothly from planning to implementation?
- How could DMPs usefully support the work, and what specific functionalities of a DMP tool would help?

[1] For example Bakos, Asztrik; Miksa, Tomasz; Rauber, Andreas (2018). Further development in RDA group https://www.rd-alliance.org/madmps

- What ideas (solutions, tools, resources) can be helpful for a successful implementation of the most desired functionalities of DMPs?

Sets of positive and negative scenarios were collected and further analyzed, resulting in a collection of ideas to be further developed during the writing process and feedback loops for this paper.

What is a DMP and what is not?

During our first workshop on best practices for DMP templates, a fundamental question was raised about the relationship between DMPs and data documentation. Both contain metadata about the research project's data, but they operate on different levels. It therefore seems necessary to define the specific scope for each.

DMPs should include **information on how data will be handled and documented throughout the project and how they may be reused afterwards**. They guide stakeholders not only to the actual data documentation, such as in a local electronic lab notebook, but also to the procedures and standards followed by the project team during the research. In addition, DMPs document information about the long-term fate of the data after the project is completed, covering aspects such as archiving requirements and planning for data reuse. In this sense, they **contain metadata about data documentation**. Data documentation, on the other hand, starts with the collection of research data and may continue until the end of the project, or, if data are still curated afterwards for the long-term, it may be an ongoing, never-ending process. It follows predefined rules, vocabularies or standards set out in the DMP for all metadata types listed there. A DMP cannot therefore replace data documentation, but it does define the key points and steps involved.

In addition, understanding what a DMP is not can help clarify its purpose and scope. A DMP is **not a guideline on how to conduct research with data**. While a research guideline provides information for a comprehensive understanding of the research process, a DMP addresses only the data-related aspects during the research process and guides the responsible handling of research data. A research guideline provides information on how to conduct research and offers detailed instructions and recommendations on, for example, hypothesis formulation, analysis techniques, and interpretation of results. In contrast, a DMP focuses on the generation, organization, documentation, preservation, and publication of research data. It outlines data formats, data reuse, data storage, and backup procedures, as well as responsibilities, costs, and resources. Additionally, it ensures ethical, legally compliant, and secure handling of the data. A DMP explains how data will be managed at different stages of the research project, from data collection to archiving and possible future use. It is part of the project management. Both entities of the research process are interrelated and serve their specific purpose: research guidelines as part of research methodology and DMP as part of data management.

Although a DMP is often submitted as **part of a research proposal**, it is distinct from the proposal itself. A research proposal outlines the research project's aims, objectives, methodology, and significance of the research project, whereas the DMP primarily addresses data-related aspects.

A DMP is **not a data analysis plan, nor is it a research data policy**. It does not detail the specific analyses or experiments that will be conducted with the data. Instead, it focuses on data management practices. Data policies govern how research data should be managed. In Germany, there are almost no policies with detailed requirements, but mostly only basic commitments to promote responsible and transparent research data practices. In general, a DMP does not provide instructions for action like a policy does, but it is a necessary building block or, respectively, an instrument for the implementation of the policy instructions.

There is no rigid structure to the DMP due to the large number of possible tasks. It is therefore **not a static document**. A DMP is a living document that can evolve over the course of a research project. It can be updated as data management needs change, for example when new data sources are discovered, or when data sharing plans are revised. To date, DMPs are **not universally standardized**: While there are common elements, generic templates and best practices for creating a DMP, the specific requirements, components and formats may vary depending on the purpose and benefits sought by the respective users, funding agencies or institutions. Users should tailor their DMP to their specific needs, which means that DMPs are **not just for scientific research**. Although DMPs are often associated with scientific research, they can in fact be used in a variety of fields.

Our vision for DMPs in the NFDI

DMPs have the potential to connect different actors and target groups in the RDM process and these stakeholders have different requirements for DMPs, templates, and tools.

We discussed ideas that would extend the functionality of DMPs, requiring the integration of specific modules, e.g. for technical workflows or approaches to data encoding and publication, or that could help with strategic decision making (monitoring) and assessment (evaluation). All these ideas can be brought together by specific objectives shared by one or more of the involved groups in the RDM process: researchers, RDM consultants, data stewards, service providers, research funding support, legal advisors, senior management, or NFDI base services. Overall, we identified three interrelated leitmotifs for future DMPs: **automation, awareness, and usability**.

DMP templates should support both researchers and data providers in their planning process by referring to current best practices, standards, or specific data infrastructure requirements (awareness). They should be standardized and flexible to support different types of research, and be adaptable to specific community needs (usability). In addition, DMP tools should be interoperable with other services and tools (automation). The latter

provides the underlying structures needed to improve awareness and usability measures. In the following chapters, the collected ideas for DMPs are illustrated in more detail.

Support communication between different stakeholders

A DMP has the ability to extend the communication beyond the internal scope of a project to the different stakeholders. This will **facilitate the actual communication between the involved stakeholders** to promote a more seamless integration and flow of information across different platforms, projects, and actors. To achieve this, key considerations for implementation are outlined below.

There is a need to ensure **the delivery of information to relevant stakeholders** in order for them to receive automated updates relevant to a DMP, such as statistics on storage space and resource requirements, to ensure timeliness and accuracy. Ideally, this can be specified in the DMP, depending on which stakeholders are to receive information or take action. DMP tools can provide collaborative access for various team members, RDM staff, and others in order to provide the information needed. This includes change tracking and personalized views to highlight information that is important to different roles, while keeping track of who made what changes. This means that contacts can be automatically notified of changes, or researchers can actively request help with specific questions and contacts are notified accordingly. A mechanism such as **change notifications for key stakeholders** will be useful in this case. Automatic alerts or the ability to request help must be in place to ensure smooth communication (see also the chapter on *support strategic management*).

For **collaborative access by stakeholders outside a project**, the use as such needs to be functionally extended to manage different access rights for the respective user groups. A DMP is therefore extended from internal use within a research project to use as a project and information management tool involving different external stakeholders. In this form, it could act as a **communication platform** and provide versioning and approval processes (see also the chapter on *integration into the research process*).

DMP tools could be technically enabled to **highlight views for different audiences**, e.g. a data protection officer sees information on data protection directly, but could also access other aspects if access is deemed necessary and granted. Prior to any project a **content analysis of the DMP template** to be used must be carried out to understand what DMP content is essential for stakeholders outside the project, ensuring the appropriateness of the **information shared**. Within such a DMP, explicit timelines for (data) activities allow stakeholders to track progress and align efforts, while regular milestones and deadlines promote communication about successes, challenges and potential schedule adjustments (see also chapter *support strategic management*). However, care must be taken to avoid excessive oversight, as the need to conform to the control or external workflow expectations of more senior stakeholders, or those with authority but lacking expertise and responsibility, can lead to a loss of free (research) activity by potentially stifling innovation and independent inquiry.

Finally, the following examples illustrate other **possible scenarios for smooth and meaningful communication** of stakeholders via a DMP.

For example at the project level **between the project team and the Principal Investigator (PI)**: Collaboration between the research team and the PI is essential for the success of the research project. The role of a PI in a project is multifaceted and mainly involves scientific leadership, supervision, team coordination, communication, quality assurance, and mentoring. The success of a research project is often inextricably linked to the PI's ability to effectively manage and delegate these responsibilities to the team. The PI ensures alignment with the research objectives by reviewing and approving the DMP. For example, if a project member leaves, the project team can communicate with the PI through the DMP, outlining revised responsibilities and documenting the transfer of knowledge to ensure a seamless continuation of data management tasks (see also the chapter on *integration into the research process*).

In addition, communication via DMP should also be able to draw wider circles. For example, **between research funders and the project team**: When it comes to communication between these two stakeholders, a DMP can demonstrate accountability in data handling during project evaluations, promote transparency and adherence to established standards, provide funders with key insights into how resources are being used, and ensure alignment with funding requirements and expectations. For example, funders may request a report on the use of funds for data management, and the project team can use the DMP's guidelines to produce a report that builds trust; or, if the project team updates the DMP due to changing priorities, funders can review it to ensure alignment with new objectives, demonstrating adaptability while maintaining accountability.

And as a final example, a possible communication **between a project team and a data center or data provider** is outlined. The project team uses the DMP to coordinate with data centers to determine data storage and preservation strategies, while specifying procedures, formats, and access requirements. Ongoing discussions promote standardized data practices and long-term sustainability. For example, when archiving a dataset, the project team refers to the DMP for the necessary procedures and communicates with the data centers as outlined in the DMP to ensure accurate and efficient archiving that meets both the project's needs and the center's standards.

Making DMPs easier to complete

DMPs accompany a project throughout the research process. To increase the usability of this tool, it should be easy to create and understand for all stakeholders. Three starting points to facilitate the creation of a DMP are: 1. a clearer structure, 2. the automation of certain processes, and 3. the connection of additional data management services.

To improve the usability, DMPs should be concise, follow a **consistent structure** with appropriate response options wherever possible to be themselves interoperable and reusable, have **clear instructions** for completion, and should be stored with the research

data to be themselves findable and accessible. The questions themselves should be based on project-specific requirements. Questions on general standard services should be **prefilled with suitable options**, and **adaptations** to project- or domain-specific requirements should be possible. In addition, aspects that result directly from the answered questions and the data involved should be created automatically. Examples of **automation** would be the estimation of data management costs based on the description of the data (quantities) to be generated, the suggestion of appropriate repositories for the type and quantity of the described data, or the creation of metadata records for inserted files. In addition, DMP tools should automatically **check for completeness and quality** and, in complex subject areas such as legal issues, identify when the standard questions are not sufficient and a reference to external tools and expertise is required. Furthermore, the connection of other data management services, e.g. local services of an institute or services provided via the NFDIs, could enable the **transfer, exchange, and reuse of metadata**.

Integration into the research process

DMPs combine and contain metadata across the entire research process, thus providing the opportunity to act as a **communication hub** between researchers and RDM services. Linking DMPs to researchers' everyday tools, such as electronic laboratory notebooks (ELNs), is crucial to strengthening their integration into the research process. This connectivity will increase collaboration and efficiency in RDM by ensuring that research materials and DMPs remain linked throughout the research process, making the DMP a living document. Seamless tool connectivity is achieved by referring to a central vocabulary for RDM processes and DMP providing the necessary content and interfaces for other RDM relevant services as machine-actionable DMP. The document-centric view of DMPs needs to be replaced by a view of DMPs as a **structured metadata collection that acts as an exchange platform**.[2] It allows data infrastructure and services to communicate recommendations and requirements to the users via a DMP tool, allowing for a smooth processing of collected data in later stages of the research process and leaving the document as just one possible (and mostly not so important) manifestation or representation of a DMP.

Taking an approach that integrates DMPs into research processes and services holds potential advantages such as **enabling infrastructure providers (data curators), data centers, and local infrastructure to register services** within the DMP tool. This allows specification of necessary metadata and prerequisites via API to support automated collection when users select options during planning. This can lead to automatically created tasks for project management or checks for completeness, furthermore support for cost calculations or reminder functions. For this, DMP tool providers need to offer open and expandable DMP templates, raising awareness and enable selection of registered RDM services, which can be described with a standardized RDM knowledge graph.

[2] Bakos, Asztrik; Miksa, Tomasz; Rauber, Andreas (2018).

Simplification of the application and the review process

Another potential function of DMPs which has so far been underestimated is to support, simplify and accelerate the application and review processes. In the pre-funding phase, data stewards and RDM consultants or participating data centers could be given access to (parts of) the DMP to ensure that the described quality measures are sufficient for the deposit of datasets and comply with internal regulations. Furthermore, these checks can be automated, when data centers, infrastructure providers or even research communities register their requirements in the DMP tool (see also chapter *integration into the research process*), thus relieving RDM consultants from the need to research detailed requirements for each specific use case.

Defining these requirements in a data-style approach is therefore an important prerequisite for further automation. Some research communities have already defined specific requirements for data management in document form.[3] These have to be transformed into **community specific data fields and option sets for DMP tools** and ideally provided and updated by these communities themselves. In this way, the most current, discipline-specific and nuanced data and/or method-specific requirements are seamlessly presented to researchers. This requires a standardized structure for DMP modules.

Finally, by fulfilling the idea of being a living document, DMPs could be used as part of an **automated quality assessment**. They contain metadata along the data life cycle and the status of a project. For example, information on published or archived datasets can be linked to the holding data infrastructure via persistent identifiers (PIDs). This enables automated evaluation by funding agencies with access to the DMP, as well as by data stewards collecting information on existing datasets for reuse by others.

Support strategic management

Another desirable function of DMPs is their use for strategic management tasks. In order to **achieve strategic goals**, the special area of documentation, which is closely related to the tasks of statistical evaluation and reporting, is indispensable. With the support of DMPs, project objectives, project progress and data status can be recorded at certain stages of the project. Coupling DMPs with research information systems, e.g. based on the KDSF[4], can contribute to greater efficiency in research project management, including RDM, for all stakeholders involved as well as to far-sighted strategic planning. The data in a DMP can be used to generate regular reports on planned and ongoing projects and their needs, as well as to notify specific needs in the overall project framework (notification to RDM services, transparency of communication). In this way, a DMP collection enables a top-down view of one's own institution or a specific research area. With this information, different strategic fields of action can be better understood and, above all, changes can be responded to more quickly in a volatile world. Data can be used to inform both

[3] e.g. recommendations published by the DFG: Subject-specific Recommendations on the Handling of Research Data (2023).
[4] KDSF: Der KDSF – Standard für Forschungsinformationen in Deutschland https://www.kerndatensatz-forschung.de/.

strategic (e.g. setting research priorities) and operational (e.g. resource allocation) decisions, such as optimizing resource allocation, identifying research priorities, enhancing competitiveness, fostering interdisciplinary collaboration, optimizing research processes, improving data transparency and integrity, complying with grant requirements and reporting, and ensuring long-term data archiving and availability.

A DMP that guides researchers in complying with the FAIR Data Principles[5], data policies of third-party funders and scientific societies, and provides needs-based solution options, will greatly simplify their compliance. It will also improve the transparency and integrity of the data and increase the competitiveness of the research institution by making it easier for researchers to meet funding requirements. For planning and further developing IT services, the information in DMPs can help to calculate the current and anticipated IT resource requirements, such as storage space, data curation support, RDM-consulting or high-performance computing, in order to optimize resource allocation. This also enables better support of research processes through on-demand service activities. The information in a DMP can be also used to determine how intensively the current research priorities are being researched. It can also be used to identify trends in the development of new research priorities. Details of collaborators and the types of data and methods used can help to identify and promote opportunities for collaborative and interdisciplinary research. All of this helps to keep the institutional research strategy agile in the context of changing circumstances.

This includes, for example, the **management and oversight of large projects** and their sub-projects through a common project standard. In addition, the DMP of the sub-projects directly provides/displays what the defined data policy standards of the project network are. A DMP makes it easier to identify possible synergy effects in the distribution of tasks (**reminder function, success monitoring**). At the same time, users of DMPs have the opportunity to become **aware of prioritizing tasks** that can be completed within a realistic scope and effort. For example, the automatic transmission of metadata from a DMP to the relevant stakeholders could simultaneously generate a to-do list that also automatically adapts to changes in the project status. In this way, administrative workflows can be simplified and (partially) automated.

When it comes to **complying with specific policies and guidelines**, it may be beneficial for an institution to use DMPs to develop best practice implementation strategies, establish RDM standards and provide optimized local demand-driven RDM services in order to continuously adapt data policies and services over time. Again, it would be valuable to systematically draw on and evaluate existing DMPs. In this way, DMPs can be used as a resource management tool. Last but not least, it can also be mentioned that information from a DMP is an important component or essential criterion in the **evaluation of applications** for third-party funding.

[5] Wilkinson, Mark D. et al.: The FAIR Guiding Principles for scientific data management and stewardship (2016).

Possible next steps

DMPs play a crucial role in ensuring that all types of data contributed to the NFDI are well organized, of high-quality, and accessible to data users. They provide a structured framework for researchers and data collectors to manage their data effectively and contribute to the success of the NFDI in its mission to provide a consolidated, comprehensive and accessible research data infrastructure across all disciplines. The NFDI can benefit from all the information that is stored in DMPs in many ways, but particularly in the areas of data quality, data interoperability/exchangeability, data discovery, and data standards.

The current trend is still to tailor DMPs to the needs of specific subject communities and to adapt them to suit typical research projects in those subjects. However, the NFDI should go beyond this to identify specific use cases for research and data collection from different domains across disciplines. Such scenarios will determine the desired data product and can be used as a basis for building customized DMPs as a service. This process will require the NFDI to engage in an intensive dialogue with both the scientific and GLAM communities. Appropriate DMPs could be developed based on the desired data use or data product to be achieved. These DMP templates should then be tested by the communities and stakeholders. In a second step, their feedback can be incorporated into the revision.

User stories can be helpful in formulating the desired data use scenarios, data plan and end product. Both heterogeneous data from completed research projects and homogeneous data collections that are maintained on an ongoing basis should be considered. In order to use a DMP as a guide for a specific goal, it should not just consist of questions and open-ended answers. For many users, this structure is not concrete enough and they feel let down when planning their projects and digitization plans. At the very least, advice and implementation steps should be recommended. It would be even better if appropriate choices could be provided along the specified discipline(s), methods, and research data types to help identify the appropriate RDM services more quickly. The various guidelines and services provided by the NFDI consortia, including Base4NFDI in the areas of digitization, data enrichment and cleaning, FAIR data, legal and ethical issues, publication and preservation, provide a solid foundation as potential modules. The collected visions and ideas in this paper can also be a good starting point to address data management gaps and needs.

We also see a high need for standardization in the context of DMPs, a prerequisite for all visions presented so far. The standardization process needs to be considered at different levels.

1. In order to foster the communication between the different stakeholders involved in the data management process, a **common set of general topics and questions** should be defined for all DMPs in the NFDI at a high, first level of granularity. This would allow specific topics to be assigned to specific stakeholders. In addition, a clear structure allows

plug-in modules for domain-specific needs in data management planning for specific communities, which is particularly important for cross-disciplinary research. Based on this overall structure, consortia can collaboratively define the granularity for each topic and standardize responses and formats.

2. At a second (or also third) level of granularity consortia should define additional questions and answer sets for specific needs. **Standardized answer sets** are needed to automate certain processes. These modules should also be based on a certain set of quality criteria, defined and established by a cross-consortia process.

3. On a technical level, these questions and answers should be based on a **common vocabulary**, e.g. similar to the domain defined in RDMO. This **metadata schema** is needed for interoperability and for linking different services and tools. Questions defined in (1.) and (2.) should refer to this common vocabulary. The metadata schema is required to connect DMP tools to other NFDI services.

4. Additional **technical interfaces** may be required to support the exchange of information between services. Preliminary work, such as the maDMP or RO-crate, should be leveraged.

5. An **editorial process** and authority to maintain the common DMP structure and vocabulary needs to be implemented.

An agreement on a common DMP structure and vocabulary at the first level of granularity is urgently needed, before consortia can start to develop specific modules. Some approaches, such as the core requirements for DMPs defined by Science Europe[6], can support a rapid decision-making process. Subsequently, modules on a lower level of granularity can be defined and regularly updated by the consortia. In many aspects, these modules can be defined collaboratively by a group of concerned consortia according to the specific needs of their respective communities.

Conclusion

This position paper outlines visions for realizing the untapped potential of DMPs for the NFDI. By promoting machine-actionable DMPs and establishing a common structure and basic vocabulary, all consortia will be enabled to provide better and more integrated DMP services. In addition, the DMP can act as a communication hub between NFDI services making it a cornerstone of the upcoming NFDI service architecture.

Our aim is to connect the different stakeholders involved in the RDM process via the DMP to improve the support structure for researchers. Following the three leitmotifs of awareness, usability and automation mentioned at the beginning of this paper, there is a need to better integrate the DMP into the research process to facilitate planning and support, but also to uncover additional potential for application and strategic

[6] Science Europe: Practical Guide to the International Alignment of Research Data Management - Extended Edition (2021)

management processes. A generic framework with flexible modular extensions is the most appropriate way to develop a DMP standard. Improving the usability and integration of services into the data management planning process requires standardization at multiple levels. This will require continuous exchange between consortia on vocabulary, technical standards and DMP modules in infra-dmp as well as with other sections of the NFDI.

Acknowledgements

We thank all participants of the infra-dmp workshops and especially Sarah Wettermann, Yasmin Demerdash and Ahmed Mansour for their feedback and corrections.

Signed

Name	Institution	NFDI consortia	Email
Katja Diederichs Birte Lindstädt	ZB MED Informationszentrum Lebenswissenschaften	NFDI4Microbiota, NFDI4Health, FAIRagro	diederichs@zbmed.de lindstaedt@zbmed.de
Celia Krause	Deutsches Dokumentationszentrum für Kunstgeschichte - Bildarchiv Foto Marburg	NFDI4Culture	celia.krause@fotomarburg.de
Marina Lemaire	Servicezentrum eSciences der Universität Trier	NFDI4Memory	marina.lemaire@uni-trier.de
Marco Reidelbach	Zuse Institut Berlin	MaRDI	reidelbach@zib.de
Jürgen Windeck Gerald Jagusch	TU Darmstadt	NFDI4Ing	Juergen.Windeck@tu-darmstadt.de
Sarah Wettermann	JGU Mainz	SFB1551, FDM Manager	sawetter@uni-mainz.de
Yasmin Demerdash	IMB Mainz	IMB data manager	y.demerdash@imb-mainz.de
Ahmed Mansour	*Physics Department and CSMB, Humboldt-Universität zu Berlin, Germany*	FAIRmat	amansour@physik.hu-berlin.de
Christian Bracht	Deutsches Dokumentationszentrum für	NFDI4Culture	bracht@fotomarburg.de

	Kunstgeschichte - Bildarchiv Foto Marburg		
Sonja Schimmler	Fraunhofer FOKUS, TU Berlin	Sektion Infra, NFDI4DS, NFDI4Cat, Base4NFDI	sonja.schimmler@fokus.fraunhofer.de
Michael Diepenbroek		Sektion Infra, NFDI4Biodiversity	mdiepenbroek@pangaea.de
Petra Ritter	Charité Universitätsmedizin Berlin	NFDI4Health	petra.ritter@charite.de
Daniela Hausen	RWTH Aachen University	NFDI4Chem	hausen@ub.rwth-aachen.de
Lukas Weimer	Niedersächsische Staats- und Universitätsbibliothek Göttingen	Text+	office@text-plus.org
Martin Spenger	Universitätsbibliothek der LMU München	Research Data Management Division	fdm@ub.uni-muenchen.de
Hela Mehrtens Christin Henzen	GEOMAR Helmholtz-Zentrum für Ozeanforschung Kiel Professur für Geoinformatik, Technische Universität Dresden	NFDI4Earth	hmehrtens@geomar.de christin.henzen@tu-dresden.de
Jimena Linares	Gesellschaft für Biologischen Daten - GFBio e.V.	NFDI4Biodiversity	jlinares@gfbio.org
Ulrich Krieger	Universitätsbibliothek Mannheim	BERD@NFDI	ulrich.krieger@uni-mannheim.de
Fabian Cremer	Leibniz-Institut für Europäische Geschichte (IEG)	NFDI4Memory	cremer@ieg-mainz.de
Xiaoran Zhou	Forschungszentrum Jülich	DataPLANT	x.zhou@fz-juelich.de
Dr Karsten Tabelow	Weierstrass Institute for Applied Analytics and Stochastics (WIAS)	MaRDI	tabelow@mardi4nfdi.de

References

Bakos, A., Miksa, T., Rauber, A. (2018). Research Data Preservation Using Process Engines and Machine-Actionable Data Management Plans. In: Méndez, E., Crestani, F., Ribeiro, C., David, G., Lopes, J. (eds) Digital Libraries for Open Knowledge. TPDL 2018. Lecture Notes in Computer Science, vol 11057. Springer, Cham. https://doi.org/10.1007/978-3-030-00066-0_6.

DFG (2023): Subject-specific Recommendations on the Handling of Research Data. Last Updated 15. June 2023. available online https://www.dfg.de/en/research_funding/principles_dfg_funding/research_data/recommendations/index.html.

KDSF: Der KDSF – Standard für Forschungsinformationen in Deutschland. Available online https://www.kerndatensatz-forschung.de.

Science Europe. (2021). Practical Guide to the International Alignment of Research Data Management - Extended Edition. https://doi.org/10.5281/zenodo.4915862

Wilkinson, Mark D.; Dumontier, Michel; Aalbersberg, I. Jsbrand Jan; Appleton, Gabrielle; Axton, Myles; Baak, Arie et al. (2016): The FAIR Guiding Principles for scientific data management and stewardship. In: Scientific data 3, S. 160018. https://doi.org/10.1038/sdata.2016.18.

Abbreviations

DMP (pl. DMPs) = Data Management Plan

FAIR = Findable, Accessible, Interoperable and Reusable
FAIR data are optimally prepared and accessible to people and machines. The application of the so-called FAIR Data Principles is intended to improve the reusability of data sets. The FAIR principles were included by the European Commission in the EU Horizon 2020 funding guidelines. The NFDI is also committed to establishing research data management according to the FAIR Data Principles.

GLAM = Galleries, Libraries, Archives, Museums
The acronym GLAM refers to cultural institutions whose mission is to provide access to knowledge. GLAMs collect and care for cultural heritage materials in the public interest. As collecting institutions, GLAMs preserve and make accessible primary sources of value to researchers.

infra-dmp = Working group on data management planning of the NFDI section Common Infrastructures
The working group connects all consortia working on DMPs in the NFDI.

KDSF = Kerndatensatz Forschung
The KDSF is a metadata standard in the German science system. It describes what information universities, universities of applied sciences, non-university research institutions and other research organizations should provide about their research activities.

NFDI = National Research Data Infrastructure
The goal of the NFDI is to make data from science and research available, network it and thus make it usable in the long term. Institutions from various fields are involved in the NFDI. They are working on services, training for researchers and standards for handling data. The data will be systematically indexed for the entire German science system, networked and made usable in a sustainable and qualitative way in the form of a knowledge graph.

PID (pl. PIDs) = Persistent Identifier

RDA = Research Data Alliance
The RDA is a global community-driven organization that aims to facilitate the sharing and exchange of data among researchers and professionals in various disciplines, domains, and countries. It was established in 2013 to address the challenges related to data sharing, interoperability, and access in research.

RDM = Research Data Management

RDMO = Research Data Management Organiser

The RDMO software enables scientists to create DMPs, work cooperatively on data management projects, involve stakeholders and IT, and structure the workflow thanks to created tasks. RDMO can be customized. This creates needs-based questionnaires for various use cases and disciplines. Many facilities in Germany have already installed RDMO.