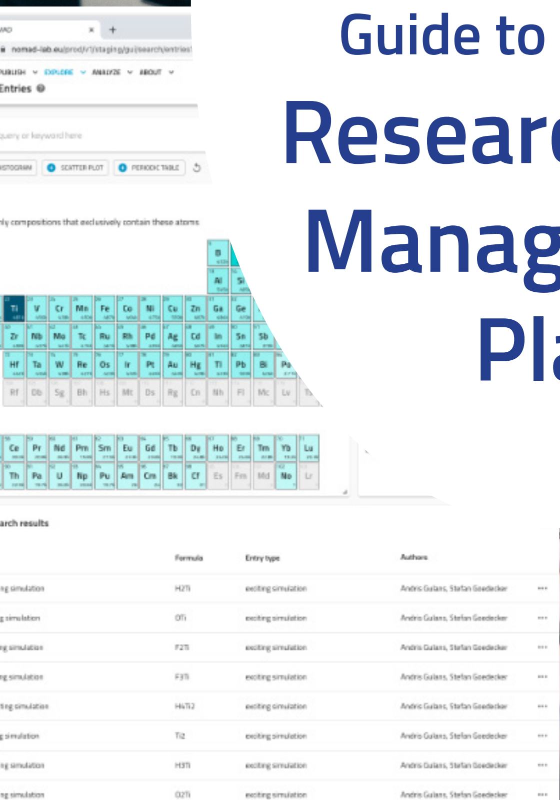
FAIRmat

# Guide to writing a Research Data Management Plan



This guide was created by FAIRmat. Cite it as “FAIRmat, Guide to Writing a Research Data Management Plan”, version 1.0, 25 March, 2023.



DOI: 10.5281/zenodo.7936477

# Contents

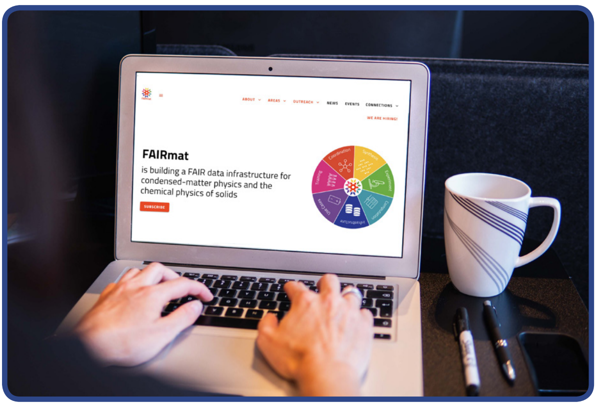
# Introduction

Welcome to this practical guide on the preparation of data management plans (DMP). If you are a researcher planning to submit a proposal to the the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation), it is important to know that all research proposals - for both individual and collaborative projects - should include a detailed specification on handling research data.[1]

DFG Deutsche Forschungsgemeinschaft

The DFG requirements are based on a catalogue of questions to cover key aspects of handling research data and are categorized into several sections: data description, documentation and data quality, storage and technical archiving of the project, legal obligations and conditions, data exchange and long-term data accessibility, and responsibilities and resources.[2]

A DMP is a critical document that outlines the strategy, procedure, and tools that will be used to handle research data throughout the data life cycle (shown in Figure 1). This includes the various stages of data management, starting from project initiation to project closure, and also addresses the long-term fate of the data. The primary purpose of a DMP is to ensure that data are properly described, documented, acquired at high quality standards,




Original photograph: Myriam Jessier on Unsplash stored, and archived in a way that ensures long-term accessibility and usability by the wider research community. It is a tool to improve the research process and outcomes.

A meaningful DMP goes beyond the minimum requirements of research funding organizations, and puts focus on the research data as a main outcome of the research process, rather than a secondary outcome which only serves to enable publications, dissertations, or patents.

In this guide, we will provide you with comprehensive information and practical tips specific to the fields of condensed-matter physics and materials science on creating a DMP that meets the DFG requirements and aligns your research with the FAIR data principles,[3,4] the DFG code of conduct,[5] and the EU open science policy.[6]


Figure 1: Data life cycle in research projects.

We will take you through the essential components of a DMP, provide tips on data management best practices, and offer guidance on appropriate tools and technologies. By the end of this guide, you will have the knowledge and tools necessary to create a thorough and effective DMP that not only meets the requirements of the DFG but also supports the long-term success of your research project. This guide follows the sections in the DFG checklist, and the questions to be addressed as required by the DFG are listed after each section.[2]

Although this guide is based on the requirements of the DFG, the information is not exclusive to any specific funding agency and can be used as a general guide for other research areas. Before preparing a DMP for your project, make sure to check the specific requirements for your funding agency, discipline, and research institution.

## Step 1: Determine the funding body requirements

The first step in preparing a DMP is to determine the research funding body requirements. It is crucial to keep in mind that these requirements are updated frequently, and it is recommended that you check them every time you prepare a DMP, even if you have applied to the same funding body before.

## Step 2: Consider discipline-specific requirements

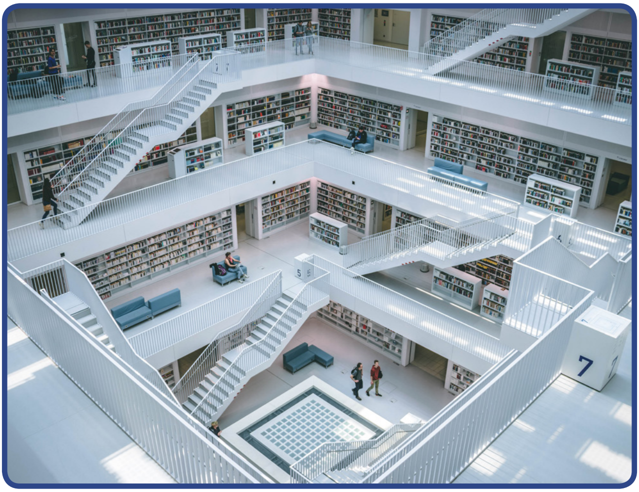
Different research communities have their own standards for how data should be managed in the respective field. Therefore, research data should be handled in a way that is appropriate to the specific discipline. It is important to be aware of any subject-specific requirements when preparing your DMP. For subject-specific and up-to-date advice on DMP preparation, we recommend contacting the relevant NFDI consortium. You can find a list of all consortia and their contact information at [www.nfdi.de](http://www.nfdi.de).

nfdi
Nationale Forschungsdaten Infrastruktur

nfdi.de

## Step 3: Seek support at your research institution

Determine any research data management requirements set by your research institute. Such information, as well as advice on DMP preparation can, for instance, also be obtained from the university's libraries or computing and data centers. They can provide valuable support and resources for researchers, and they can offer guidance on best practices for managing research data.



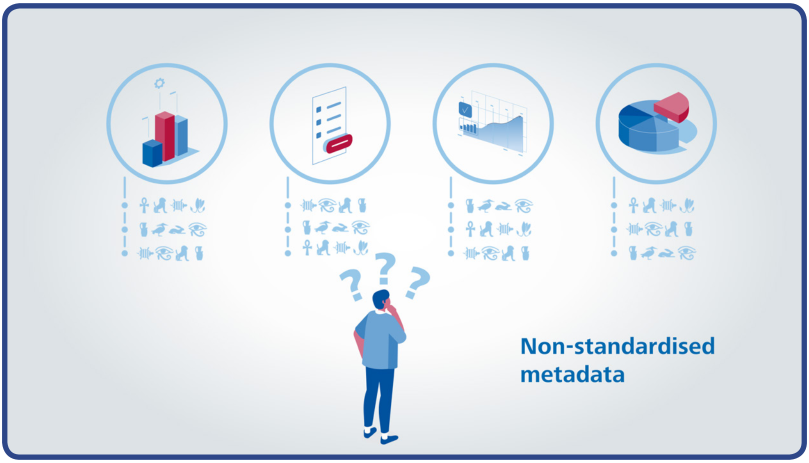
Gabriel Sollman on Unsplash

Preparing a DMP is a crucial aspect of research proposal development. By following recommendations and steps in this guide, and seeking advice from relevant resources, you can ensure that your DMP meets the requirements of your research sponsor and aligns with discipline-specific standards.

# Data Description

## Know what you need to manage

Before any management plan can be created, what needs to be managed should be clearly defined. In this section, provide a comprehensive description of the data that will be either generated by the proposed project or reused from another source. By doing this, researchers can ensure that they have a clear understanding of the data types and sources, allowing for effective management and organization of the data throughout the research process.




nfdi.de

We recommend structuring this part into two sections: data generated by the current project and data obtained and reused from an external source. A comprehensive description of data should include data type, source, format, and quantity.

## Data types

Types relevant to the fields of condensed-matter physics and materials science and compliant with the DFG categorization of data types include:

- Measurement data: Data obtained through measurement via various instruments
- Visual information: Images and video recordings
- Methodological test procedure: Protocols to prepare samples, record observations or perform an experiment
- Computations and simulations
- Objects/Samples: From an external source or prepared in the lab
- Software.

## Data sources

These can be a research instrument, methodology, software, or code that will be used to generate the data, as well as external sources from which data are obtained for reuse.

**Research instruments:** When describing research instruments used to acquire data, it is recommended to provide information such as the type, model/vendor, configuration, specific information related to the source, acquisition mode, analysis workflow, and detectors when applicable. This information will help to ensure that the data can be accurately interpreted and effectively managed throughout the research process.

**Methodologies:** For methodologies describing procedures to prepare samples, record observations, or perform an experiment, it is recommended to provide the name, workflow, and reference to a standard methodology accepted and published by the community. This information will help to ensure that the methodology can be replicated and that the resulting data can be accurately interpreted.

NOMAD extracts data from 60+ different formats and codes.

**Computations and simulations:** For computations and simulations, it is recommended to specify the computer codes or software used, and cite them appropriately.

**Custom-made instrumentation or newly developed methodologies:** In the case of custom-made instrumentation or newly developed methodologies, it is important to refer to the appropriate section in the research proposal where they are described.

**External data sources:** Data repositories, publications, online databases, standardization institutes, or vendors should also be clearly described and cited. The description should include any restrictions or agreements on reusing the data.

## Data formats

Choosing the right format for the data generated by your project ensures long-term accessibility and interoperability of your data. It is important to carefully consider the data formats that will be used. In this section of your data management plan, you should provide a comprehensive description of the chosen data formats and your justification for their selection.

Proprietary file formats, which are often used by scientific instrumentation software, should be avoided as they limit data interoperability and reusability. Instead, open and non-proprietary file formats, i.e. formats that are supported by more than one vendor and can be accessed with different software systems, are recommended. In order to prevent any loss of (meta)data, it is advised to keep two copies of the data: one in the proprietary data format and one in an open data format.

In your DMP, you should list all the formats of generated data, as well as the strategy and tools you plan to use to convert proprietary data to other open data formats. You should also describe whether the conversion of the data will result in losses of (meta)data or other data generation details, and your plan to compensate for any such loss, such as adding metadata manually.

To assist you in selecting appropriate data formats, you can refer to Table 1, which provides a non-exhaustive list of general file formats and those specific to the field of physics. Your justification for the chosen formats may include factors such as expertise of staff, accessibility, or interoperability. Ultimately, the goal is to select formats that enable data sharing with others and retrieval in the future, in line with the FAIR principles.

## Data quantity

The quantity of each data type should be defined in terms of volume, such as gigabytes or terabytes, and the number of files. It is important to associate each data type with its own quantity and provide a total sum of all data in the DMP. This information will help to estimate the storage and backup requirements, as well as the potential costs associated with managing and preserving the data. Therefore, it is recommended to carefully estimate the expected quantity of each data type and to consider any potential growth or changes over time. It is also important to ensure that the storage capacity and backup systems are sufficient to accommodate the data volume, and to have a plan in place for handling any unexpected increases in data volume.

With NOMAD Oasis you can store and analyze an unlimited quantity of data.

Table 1: Non-exhaustive list of general file formats and those specific to the field of physics.[8,9]

| Category | Data type | Recommended | Non-recommended |
|---|---|---|---|
| General data formats | Text | TXT, HTML, RTF, PDF/A | DOC, PPT |
| | Tabular | CSV, TSV, SPSS portable (.por) | XLS |
| | Images | TIFF, JPEG2000, PNG | GIF, JPG |
| | Multimedia | Container: AVI, WAV, MP4, Ogg Codec: Theora, Dirac, FLAC | Windows Media Video, QuickTime, H264 |
| | Structured | XML, RDF | RDBMS |
| Formats common in physics and materials science[10] | | HDF5, NeXus, JCAMP-DX | |

## DFG Checklist

- How does your project generate new data?
- Are existing data reused?
- Which data types (in terms of data formats like image data, text data or measurement data) arise in your project and in what way are they further processed?
- To what extent do these arise or what is the anticipated data volume?

# Documentation and data quality

## Create a record of all the workflows for generating, structuring, and analyzing research data

When it comes to managing data, defining them comprehensively is just the beginning. To ensure that your data are organized, easily searchable, and of high quality, you'll need to develop a strategy for managing them effectively.

Here are some practical steps you can take:

1. Organize your data: Create a clear structure for your data that makes it easy to find and access. This may involve creating folders or using a database management system to keep your data organized.

2. Associate data to rich metadata: Metadata provide the contextualization and description of a data object, which allows for that data object to be properly interpreted and turned into information. Develop a system for creating metadata that accurately describes your data, including information such as data type, date created, and creator.

3. Document everything: Keeping detailed documentation of your data using appropriate tools is crucial for ensuring its integrity and usability. Create a system for documenting your data that includes information on how it was collected, processed, and analyzed.

4. Ensure data quality: To ensure that your data are reliable and accurate, you wil need to develop a strategy for quality assurance. This may involve using automated data checks, conducting manual reviews, or establishing data validation protocols.

The DMP should comprehensively describe the strategy and procedure related to the aspects mentioned above.

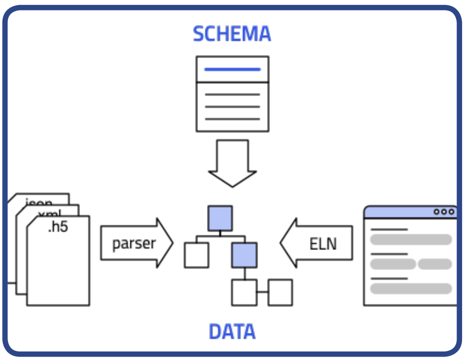
## Organizing data

If you are conducting research in the fields of condensed-matter physics and materials science, you know that the structure of any material or system is crucial to its properties. The same is true for research data. Effective data management relies on how we organize and structure our data. To keep your data organized and easily searchable, keep the mantra “structure, structure, structure” in mind.

NOMAD’s parsers automatically structure your data into hierarchically organized and connected data.

There is no one-size-fits-all approach to structuring research data due to the wide range of project scopes and research activities in the field of materials science. However, there are some general principles that can guide your efforts. You might consider structuring your research data by project phase, instrumentation used, materials properties, or any other relevant aspect.

However you choose to structure your data, it is essential to establish file naming conventions that are concise, descriptive, and easy to understand by both humans and machines. This helps to ensure that your data are easily identifiable and interpretable, both for yourself and for other researchers who may use or build on your work. For help with this, you can refer to RDMkit[11] for useful tips and examples of effective file naming conventions.




NOMAD makes it easy to structure and organize (meta)data.

By taking the time to structure and organize your research data, you will save yourself time and frustration in the long run, and make it easier for yourself and others to build on your work.

## Associate data to rich metadata

To ensure that your research data are easily accessible and usable by both humans and machines, it is essential to associate all data with metadata. Metadata should be prepared and included in your documentation, using widely accepted ontologies and vocabularies within your community.

NOMAD Metainfo defines a general domain-independent super structure, as well as highly detailed, specialized data from specific methods, tools, and programs.

The minimum set of metadata needed to characterize data varies depending on the domain, and in the case of materials science, there are many sub-domains. The first branch point is between theoretical and experimental research, with sample synthesis being partly (implicitly) included in the latter.

For experimental research, common metadata for sample synthesis and preparation should include information about all actions taken during preparation, including ingredients, conditions, durations, configurations of apparatus, as well as settings and types of instruments used.

Once samples are prepared, they are typically subject to measurement using instruments. Typical metadata for measurements should include the measurement type, instrument type, instrument configuration, instrument settings, time and duration of measurement, and physical properties measured. It is important to keep track of the sample's condition during its lifetime and to know at what point in the sample's history a particular piece of primary data was measured.

Measurement data can be attributed to a specific point in a sample's history, and this data subsequently goes through post-measurement analysis steps. These steps can include data processing, data compression, data aggregation, comparison of data instances, and combination of data instances derived from various measurement types. The analysis phase is a crucial piece of the experimental workflow, as it transforms raw data into derived data, knowledge, and insight.

To ensure that your metadata are as rich and detailed as possible, it is important to describe the metadata standards that will be adopted in your project. You should also refer to the standards adopted for domain ontology (knowledge representation: concepts, terms, and relationships) and community-accepted vocabulary. Additionally, be sure to check for subject-specific metadata schema, as these can provide more specialized guidance on how to structure and document your data.

elenabsl at shuttershock.com

By paying close attention to your metadata and documentation, you will ensure that your research data are not only usable for yourself, but for other researchers in your field as well.

## Methods and tools for data organization and documentation

In order to ensure effective data organization and documentation, it is important to plan and implement a comprehensive strategy that includes the selection of appropriate tools and methods. Determine the best way to organize, monitor, and document your data. This may involve using (a combination of) electronic lab notebooks, online repositories, spreadsheets, and other tools.

To ensure that the data are well-documented and that the associated metadata are comprehensive, it is important to use standard formats and vocabularies that are accepted in your community. This will make it easier for others to understand and use your data and written documentation. Additionally, it is important to include information about the provenance of the data, such as who generated it, when it was generated, and how it was processed.

When selecting tools and methods for data organization and documentation, it is important to consider factors such as ease of use, scalability, and the ability to integrate with other tools and platforms. It may be useful to consult other researchers in your field to identify best practices and tools that have been successful in other projects.

Overall, effective data organization and documentation requires careful planning and attention to detail. By using appropriate tools and methods and following established best practices, you can ensure that your data are well-organized, well-documented, and easily accessible to others in your research community.

## Data Quality

Quality assurance and control (QA/QC) of research data is a crucial aspect of any research project. It ensures that the data collected is of high quality and meets the expected standards, thus improving the credibility and reliability of research findings. Therefore, it is important to describe in your DMP the procedures and tools you will employ to ensure high data quality. The DFG Code of Conduct requires that the quality assurance mechanisms used are explained when research findings are made publicly available (Guideline 7).[5]

To ensure the quality of your data, you should define the standards or procedures employed in your research to measure, assess, and improve data quality. These may include regular data review, error correction, and validation against accepted standards or guidelines. It is also important to provide training activities for graduate students, postdocs, and scientists in the project to improve data literacy, acquisition, analysis, and reporting, and to detect human bias.

Additionally, you should specify the frequency and procedure for instrument calibration, conditioning, and verification tests to ensure accurate data acquisition. You should also list the statistical and visualization approaches you will use to detect errors in the data. Furthermore, you should mention any community standards adopted in carrying out and reporting of experimental and computational outcomes.

By following these guidelines, you will ensure that your research data are of high quality and meets the expected standards, thus promoting confidence in and reproducibility of your work, and preventing the dissemination of inaccurate or false conclusions.

**DFG Checklist**

- What approaches are being taken to describe the data in a comprehensible manner (such as the use of available metadata, documentation standards or ontologies)?
- What measures are being accepted to ensure high data quality?
- Are quality controls in place and if so, how do they operate?
- Which digital methods and tools (e.g. software) are required to use the data?

# Storage and technical archiving

## For the love of data, keep what you hold in your heart safe and secure!

Storing and archiving research data in a secure and accessible way is crucial for good research practice. In this section, describe your approach to storing and archiving data and metadata, as well as the planned procedures for dealing with sensitive data during the project.

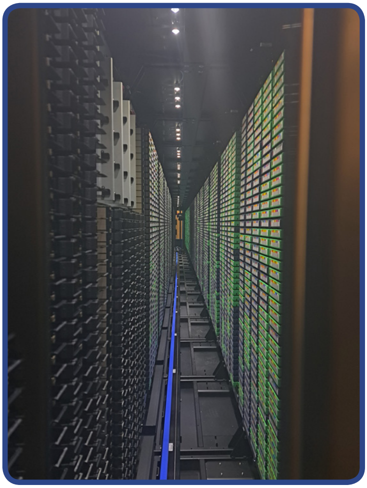


Image courtesy of MPCDF

This section should deal with the short-term storage and archiving during the project period and completion. The objective is to ensure data accessibility among the project team and external collaborators, and to prevent loss of data due to unforeseen circumstances such as hardware malfunction, theft, or natural disasters.

Start by describing the strategy for storing and backing up research data. A general recommendation is to maintain three copies of your data, using at least two different storage media. One of these copies will be stored in a physically separated site to prevent loss due to unforeseen circumstances. Decide on the frequency of creating backups according to the sensitivity of the data and the difficulty associated with their replacement in case of data loss. Include a description of the backup methods.

NOMAD adopts a conventional file system that can be easily integrated in your existing storage and backup solutions.

The backup methods can be either:

1. Full backup: Copy all of your files for every backup.
2. Incremental backup: Only copy files that have been added or changed since the last backup.
3. Differential backup: Copy any files that have been added or changed since the last full backup. Specify if the backup will be run manually or automatically (recommended), and the frequency of testing the backups.

In addition to digital data, storage of physical data, such as samples, should be considered. In this case, the storage strategy should include the physical medium and its location, as well as any storage conditions necessary to preserve the samples, such as humidity and temperature controls. For example, will the samples be stored in a nitrogen glovebox, vacuum desiccator, special encapsulation, or an ultra-high vacuum chamber? Furthermore, also describe the stability of physical samples and their deterioration over time.

Access and usage rights of the stored and archived data during the lifetime of the project should be specified, especially if the project includes several members and collaborators.

**DFG Checklist**

- How is the data to be stored and archived throughout the project curation?
- What is in place to secure sensitive data throughout the project duration (access and usage rights)?

# Legal obligations and conditions

## Do it right: legally and ethically!

It is essential to be aware of legal issues and any specific details prior to initiating the research process. Researchers should clearly describe legal issues in the DMP of their research proposal and consult their research institution's policies and requirements. We note that materials-science data are not so critical in the sense that they do not concern highly-confidential data about people (like patients in medical science). Still, the data must be handled with care.

Tingey Injury Law Firm on Unsplash

Some aspects of legal issues related to research data in materials science are discussed below.[12,13]

### Data ownership

Data ownership refers to who has the right to control access to the data generated by the project. The owner of the data could be the researchers who created the data, the principal investigator, or the research institution where the project is taking place. Researchers should explain any restrictions or conditions for accessing the data during the project and mention any contractual agreements made in collaborative projects within and across research institutions with respect to data ownership and authorship rights.

NOMAD Oasis allows you to store data privately or share it.

### Intellectual properties

Researchers should explain whether the project will generate any products that are protected by intellectual property rights, and mention which laws and policy apply. This is especially important in the case of patents, original software, or commercialized products.

All data shared publicly in NOMAD are under CC-BY-4.0.

## Copyrights

Copyright law applies to original works of authorship, including scientific publications and other research outputs. Researchers should ensure that they are complying with copyright laws when publishing or sharing their research outputs.

## Data license

Researchers should consider which data license to apply to their research data. A license can specify the terms and conditions for the use and distribution of the data. Researchers should choose a license that aligns with their data sharing objectives and complies with legal requirements.

In conclusion, researchers in the field of materials science should be aware of legal issues related to research data and ensure that they comply with relevant laws and policies. By including legal considerations in their data management plan, researchers can mitigate potential legal issues and ensure that their research data are managed appropriately.

## DFG Checklist

- What are the legal specifics associated with the handling of research data in your project?
- Do you anticipate any implications or restrictions regarding subsequent publication or accessibility?
- What is in place to consider aspects of use and copyright law as well as ownership issues?
- Are there any significant research codes or professional standards to be taken into account?

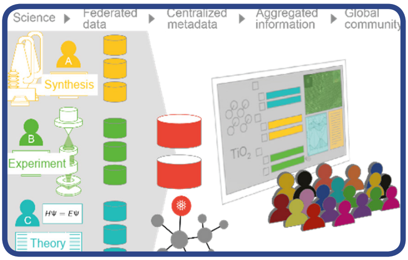
# Data exchange and long-term accessibility

## The project has ended, long live the data!

This section mainly describes the fate of the research data after the project is completed. Two key aspects should be included; data preservation and data dissemination.

### Data preservation

Data preservation is the process of securely preserving research data in a way that ensures its long-term availability and accessibility. To plan for data preservation, it is important to specify the time-frame during which the data will be preserved, which is usually decided by the requirements of the funding agency or the research institute. Good scientific practice typically entails at least ten years of data preservation.




A federated FAIR infrastructure for materials data

While researchers can manage data preservation personally, it is recommended to outsource it to institutional or third-party repositories. A repository is a platform that manages the appraisal, preservation, and accessibility of materials on a long-term or permanent basis. Data repositories are available at universities, and there are numerous field-specific or general science repositories. A comprehensive list of repositories can be found in the re3data directory.[14] To select a trustworthy repository, criteria such as offering a persistent identifier for uploaded data (such as DOI), supporting widely accepted metadata, providing a clear license and terms of use, and having a reputable hosting institute should be considered.[15] More information on selecting a repository can be found in relevant references.[14,15]

Researchers should specify the repository that will be used for preserving their data and explain how the data will be prepared to comply with the repository's requirements to support interoperability and reusability. Preparation of data for a repository includes complying with its requirements of file formats, metadata schemas, and license.

NOMAD allows you to publish and archive your data for free. Assign a DOI to uploaded datasets and reference your papers.

## Data Dissemination

Data dissemination involves making research data publicly available, either as a whole project or a subset of it. Researchers should describe in their plan whether their data will be available immediately or after an embargo period. Embargo periods are needed to protect data, such as during the writing of publications or in case of industrial collaborations where data may be kept secret for a certain amount of time.

Decide what is the right type of access for your data for example open access, registered access, with authentication procedure or controlled access, or via Data Access Committees (DACs).[16] When a dataset cannot or shouldn't be archived, a justification for such a decision should be provided. In this case, it is good practice to at least publish the metadata of your datasets.

### DFG Checklist

- Which data sets are especially suitable for use in other contexts?
- Which criteria are used to select research data to make it available for subsequent use by others? Are you planning to archive your data in a suitable infrastructure? If so, how and where?
- Are there any retention periods?
- When is the research data available for use by third parties?

# Responsibilities and resources

## Highly effective teams have clearly assigned roles, responsibilities, and resources

It is important to consider the personnel and physical resources needed to execute and control the described data management strategies, tasks, and tools. Each member or organization participating in the research proposal should be assigned a clear role based on their expertise and time availability. This ensures that everyone knows what is expected of them when dealing with research data.

To properly assign responsibilities, start by outlining the various tasks required to execute the DMP and list all the necessary skills for each task. Use this information to assign roles to existing team members, identify training needs, or determine if new hires are needed. The funding body will evaluate the submitted DMP based on the availability of qualified staff to carry it out and any proposed training plans.


Talk at the FAIRmat users meeting 2022. Photo courtesy of N. Günes.

Roles in the DMP can include data administrators, data collectors, data stewards, metadata generators, staff responsible for instrument operation and quality control, computing staff responsible for data storage and archiving, and external data centers responsible for long-term storage. However, data-management roles are not limited to the ones listed here.[17]

Make sure to include the costs and resources needed for data management in the project budget. The DFG may cover certain costs, such as those associated with preparing research data for reuse and transferring data to a public repository, to ensure data availability and reusability.[18]

### DFG Checklist

- Who is responsible for adequate handling of the research data (description of roles and responsibilities within the project)?
- Which resources (costs; time or other) are required to implement adequate handling of research data within the project?
- Who is responsible for curating the data once the project has ended?

# References

1. The Deutsche Forschungsgemeinschaft (DFG). Specification of Requirements Relating to the Handling of Research Data in Funding Proposals. *DFG Information for Researchers No. 25* (2022). https://www.dfg.de/en/research_funding/announcements_proposals/2022/info_wissenschaft_22_25/

2. The Deutsche Forschungsgemeinschaft (DFG). *Handling of Research Data: Checklist for planning and description of handling of research data in research projects* (2021). https://www.dfg.de/download/pdf/foerderung/grundlagen_dfg_foerderung/forschungsdaten/forschungsdaten_checkliste_en.pdf

3. M. D. Wilkinson et al. *The FAIR Guiding Principles for scientific data management and stewardship.* Sci Data 3, 160018 (2016). https://doi.org/10.1038/sdata.2016.18

4. M. Scheffler et al. *FAIR data enabling new horizons for materials research.* Nature 604, 635–642 (2022). https://doi.org/10.1038/s41586-022-04501-x

5. Deutsche Forschungsgemeinschaft (DFG). *Guidelines for Safeguarding Good Research Practice Code of Conduct.* https://wissenschaftliche-integritaet.de/en/.

6. European Commission. *The EU open science policy.* https://research-and-innovation.ec.europa.eu/strategy/strategy-2020-2024/our-digital-future/open-science_en.

7. K. Briney. *Data Management for Researchers: Organize, Maintain and Share your Data for Research Success.* (Exeter: Pelagic Publishing, UK, 2015).

8. Max Planck Digital Library. *Research Data Management: Data Formats. Information Platform for Max Planck Researchers* https://rdm.mpdl.mpg.de/before-research/file-formats/.

9. Humboldt-Universität zu Berlin, Computer and Media Service. *Choose data format: recommendations for the choice of a file format for long-term archiving.* https://www.cms.hu-berlin.de/en/dl-en/dataman-en/work/secure/fileformat.

10. DATACC. *Standard formats in physics. ESTABLISHING A DATA MANAGEMENT PLAN* https://www.datacc.org/en/best-practices/establishing-data-management-plan/standard-formats-in-physics/.

11. ELIXIR Research Data Management Kit (RDMkit). *What is the best way to name a file?* https://rdmkit.elixir-europe.org/data_organisation.html#what-is-the-best-way-to-name-a-file.

12. E. Böker. *Legal Issues. Forschungsdaten und Forschungsdatenmanagement* https://forschungsdaten.info/praxis-kompakt/english-pages/legal-issues/ (2022).

13. Humboldt-Universität zu Berlin - Computer and Media Service. *Share your data: Legal aspects.* https://www.cms.hu-berlin.de/en/dl-en/dataman-en/share/legal-aspects.

14. re3data.org. *Registry of Research Data Repositories.* https://doi.org/10.17616/R3D.

15. Science Europe. *Practical Guide to the International Alignment of Research Data Management* (2018).

16. ELIXIR Research Data Management Kit (RDMkit). *Data life cycle: Sharing.* https://rdmkit.elixir-europe.org/sharing#what-should-be-considered-for-data-sharing.

17. T. Valentine & DataONE. *Define roles and assign responsibilities for data management* (2011). https://dataoneorg.github.io/Education/bestpractices/define-roles-and.

18. The Deutsche Forschungsgemeinschaft (DFG). *Information on the resources available* (2022). https://www.dfg.de/en/research_funding/principles_dfg_funding/research_data/resources_available/index.html.

# About FAIRmat

FAIRmat is a consortium of the German Research-Data Infrastructure, NFDI. FAIRmat is creating a federated data infrastructure for materials data and supporting scientists in the field of condensed-matter physics and the chemical physics of solids in making their research data FAIR (Findable, Accessible, Interoperable, and Re-purposable).

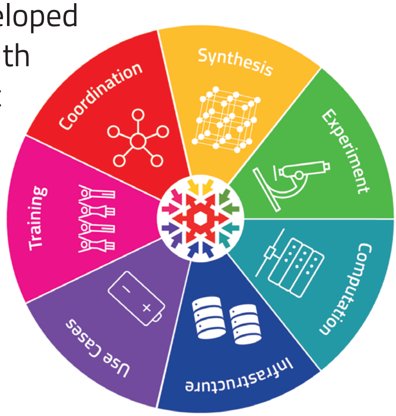
## Our mission

1. Develop and maintain a federated FAIR data infrastructure for materials data with built-in tools and standards to support scientific collaboration and proper research data management (RDM) practices.

2. Support the scientific community to introduce and maintain high standards of reproducibility, research integrity, and compliance with ethical and legal requirements by adopting proper RDM practices based on the FAIR data principles.

## How we work

FAIRmat is governed by a complementary team of internationally renowned researchers, who are actively embedded into their (sub) communities. Its infrastructure is developed by a dedicated team of coworkers with complimentary expertise. FAIRmat integrates data obtained from synthesis, experimental characterization, theory and computations, and various applications. The organization is divided into seven Areas, each addressing specific key aspects of the project.



## Our services

**NOMAD** is a free web-service that lets you share your data or use comprehensive data that others provide. You can use NOMAD to organize, analyze, share, and publish your materials-science data, as well as explore, download, and analyze your colleagues' data.

**NOMAD Oasis** lets you create your own NOMAD. Get an instant overview of all your group's data, increase productivity, and implement research data management plans with ease.

**The NOMAD Encyclopedia** allows users to see, compare, explore, and understand materials data.

**The NOMAD Artificial Intelligence Toolkit** is a platform for running (Jupyter) notebooks to analyze the data contained in the NOMAD Archive with AI tools.

**FAIRmat** offers training and educational materials on RDM, writing RDM plans, and using NOMAD tools.

# Notes

# Stay in touch!

www.fairmat-nfdi.eu

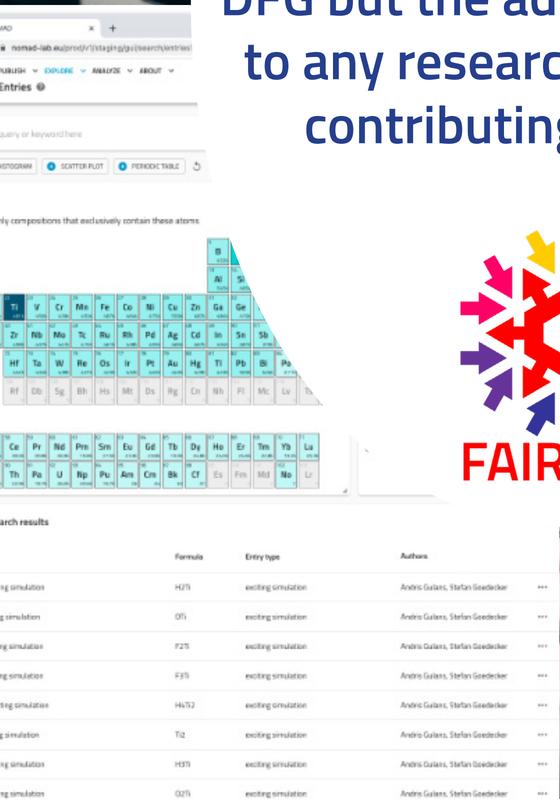
fairmat@physik.hu-berlin.de

@FAIRmat_NFDI

fairmat-nfdi

@TheNOMADLaboratory

This quick guide outlines what a Data Management Plan is, why it is important, and how to approach writing one. It includes guidelines from the DFG but the advice is relevant to any researcher writing or contributing to a DMP.




FAIRmat