Catalog NFDI4ING DMP Template (Version 2)

Comment: This is the PDF version of the NFDI4ING DMP Template, Date: 2024-09-13 Version: 2.0.0 compatible with RDMO version: 2.1.0+

To see all predefined answer options, please log in to https://rdmo.nfdi4ing.de/.

Section Introduction

Page Please add a dataset to your plan

Help: The present catalog is the checklist published by the DFG supplemented by engineering content. The questions from the DFG checklist must be taken into account when submitting an application and completing the project and should help you to structure your data management. During the interview, you will be asked several questions about the various aspects of data management. Depending on your research project, you may work with different types of data, so it might be necessary to plan different measures for each. Therefore, create as many sets as it makes sense for your approach. A logical unit for your research data could be, for example, based on

- the methodology (e.g. measurements, eye-tracking, luminance measurement, questionnaires),
- data type (e.g. images, measured data, geological data), or
- accessibility of the data (public data, data of industry partner).

You must create at least one set.

Section Data description

Page Content data description

Question: What kind of dataset is it?

Question: How is new data created in your project?

> **Help:** In engineering, the way in which data is created in your project can depend heavily on the individual sub-disciplines - such as electrical engineering, information technology, civil engineering, etc. Therefore this information is relevant for all types of data.

Question: Will existing data be reused?

> **Help:** Existing data is often reused in engineering. They can form the basis of a new project or can only be used for parts of the project. This also includes data that you use to compare your data. The NFDI4Ing Data Collections Explorer helps you to find suitable data resources. The sources of the reused data range from working group internal databases, e.g. for DFT calculations, to open and freely accessible databases such as Kaggle, to (openly accessible) repositories such as Zenodo or, more specifically, the Sequence Read Archive (SRA). Very often data from databases like Open Quantum Materials, Inorganic Crystal Structure Database, Materials Project, AtomWorkAdv., or MAGNDATA are reused.

Page Technical data description

Question: Which data types, in terms of data formats, are created in your project?

> **Help:** When choosing the file format, it is advisable to use formats that are standardized, non-proprietary and generally used in engineering. Please keep in mind that the data can be used for more than 10 years and should not lead to conflicts in collaborative use or subsequent use. When it comes to the question of origin, the initial creation of data is meant, as well as the processing and further processing of the data during the project period.

Question: How is the data processed in your project?

Question: To what extent are these incurred for your data set or what data volume can be expected for the data set?

> **Help:** Please also state the unit of measure.

Section Documentation and data quality

Page Documentation

Question: Which approaches are followed in order to describe the data in a comprehensible manner (e.g. use of existing metadata or documentation standards or ontologies)?

> **Help:** Metadata standards are used for the structured documentation of data. Here, both general standards such as ISO or Dublin Core as well as engineering-specific standards such as Metagenomics Sequences Sample (MIMS), International Data Spaces Information Model or ACM Computing Classification are used. If you are not using a metadata standard, please provide individual metadata for this question. These can be, e.g. the device or simulation settings, the test parameters and the workflow. The software (incl. version) and the hardware also count here. Administrative metadata is also often documented, such as the date, name of the responsible scientist, the project, etc.

Question: What digital methods and tools (e.g. software) are required to use the data?

> **Help:** In order to be able to reuse data, in addition to the data itself, the software, devices, etc. and knowledge of special procedures for use are required. Since use can vary greatly depending on the sub-area in engineering, please exchange ideas with your colleagues.
>
> As with the formats, the more standardized, open and established they are, the easier it is usually to reuse them. The establishment of standards for methods, the use of software, the collection of research data and the description of research results is an essential prerequisite for the comparability and transferability of research results.

Page Data Quality

Question: What measures are taken to ensure high quality of the data?

> **Help:** The quality of the data can be secured by various measures and depends on your engineering sub-discipline.

Question: Are quality controls planned and if so, how?

> **Help:** Quality controls can be done in different ways. They can be heavily dependent on your engineering sub-discipline, on your structures at the institute or in the project. It is important that the quality control is independent of the measures taken and refers to the product "data".

Section Storage and technical backup during the project

Page Data storage during the project

Question: How is the data stored and backed up during the project period?

> **Help:** During the project, different storage locations for the actively used data may be used. <u>Note</u> that depending on the storage location, your data is <u>not</u> automatically secured. You can find out which storage media have which advantages and disadvantages on www.forschungsdaten.info.

Page Technical data backup during the project

Question: How is the security of sensitive data in this dataset guaranteed during the project period (access and usage management)?

> **Help:** Sensitive data can appear in different contexts, e.g. the collection of personal data or data that arises in industrial cooperation. If you have any questions about personal data and its security, please contact the data protection officer at your university or research institution. In the case of industrial cooperation, the access and usage rights are usually recorded in the non-disclosure agreements, e.g. the implementation of a restrictive rights management. Employees are made aware of their special rights, some of which are also recorded internally in writing in the form of contracts. If you have any questions, you can contact the ELSA section, for example (for ethical, legal and social aspects) of the NFDI.

Section Data exchange and permanent accessibility of data

Page Reuse

Question: Does this data set lend itself to subsequent use in other contexts?

> **Help:** Please consider whether your data is of interest beyond your project and can be reused. Also consider how easily your data can be recreated.

> Some data can in principle be recreated at any time. Examples of this are scientific experiment data or digital copies of analog objects (as long as the originals are not lost). The effort and costs for this can be quite considerable. With regard to the question of the need for archiving that goes beyond 10 years, the effort involved in re-creating the data should be weighed against the effort and costs of archiving in these cases. Other data, on the other hand, cannot be collected again per se. This is the case, for example, with any kind of episodic observation, be it social science or natural science, as these depict a specific phenomenon at a specific point in time and/or place and are therefore usually not repeatable. Their value for subsequent use by others as well as the loss of long-term archiving that has not taken place or has failed is incomparably higher than that of reproducible data.

Question: According to which criteria is this data record selected in order to make it available for subsequent use by others?

Question: When can this data set be used by third parties?

> **Help:** The DFG Guidelines for Research Data Management Accordingly, research data should be made available as soon as possible and ideally follow the FAIR principles. The importance of data publication is underscored once again by the Guideline 13 of the Code of Good Scientific Practice. Discipline-specific practices are pointed out here. For the publication of research data, it is recommended to use subject-specific repositories.

> Recommended directories are: re3data.org and FAIRsharing.org. Currently, in addition to NOMAD and Kadi4Mat, the generic repository Zenodo is also widely used.

Page Archiving

Question: Are you planning to archive your data in a suitable infrastructure?

> **Help:** In accordance with good scientific practice, research data must generally be stored for 10 years (see Guideline 17 of the Code of Good Scientific Practice). In order for the data to remain available and findable, it must also be maintained beyond the project.

> If you archive the data in a data center or repository, the responsibility is usually assumed by the operator. Please check or clarify the responsibilities in advance. The archive service of the university/research institution can be free of charge for members and secure storage of the data is often guaranteed in accordance with the requirements of good scientific practice. Archiving includes metadata, documentation and any relevant code/software. The NFDI4Ing Data Collections Explorer helps you to find suitable data repositories

Question: If yes, how and where?

Question: Are there any blocking periods?

Page Publication restrictions

Question: Are there any impacts or limitations on subsequent publication or accessibility?

> **Help:** Please keep in mind that the data can be published after a certain blocking or embargo period and that you can also prepare this publication now. In addition to the various services in NFDI4Ing, the RDM team at your university or research institution will also support you when it comes to data publication. You can also find more information in [Guideline 13 of the code on Good Scientific Practice](). There are many reasons such as reproducibility, description of workflows and formats, added value for engineering and related sciences, FAIR data and data quality that speak for data publication. Nevertheless, there are reasons that speak against a (direct) publication of the data. This can be, for example, the protection of personal data, the commercial or industrial use of the data, or confidentiality related to security issues.

Section Legal obligations and framework conditions

Page Scientific specifics

Question: Are there important scientific codes or professional standards that should be taken into account?

> **Help:** In addition to general scientific codes, there are also engineer-specific codes, for example the „Ethical principles of the engineering profession" or the acatech-Projekt „Responsibility in the technological sciences". If you do not know these, the DFG's page on scientific integrity will help you further. Here you will find references to laws, subject-specific comments on the guidelines and explanations as well as case studies.

Page Legal Restrictions

Question: How are usage and copyright aspects as well as ownership issues taken into account?

> **Help:** With a data set, the various legal aspects must be taken into account. Whether your data is subject to copyright protection, depends - among other things - on the level of originality. Information on this can be found, for example, in the Information sheet on copyright protection of research data or at Research Data Copyright and Licensing. You can find answers to legal questions about Open Science in the book: „Rechtsfragen bei Open Science" by Kreutzer and Lahmann.

Question: What legal peculiarities exist in connection with the handling of research data in your project?

> **Help:** Examples include: Does the legal situation in different countries have to be taken into account? Does every employee have the same access rights? When collecting and processing personal data or in case of industrial cooperation additional measures may be required. For example, when collecting personal data you need to ensure formal consent. In this case you should contact the data protection officer at your university or research institution. In the case of industrial cooperation, the access and subsequent publication of data is usually restricted by non-disclosure agreements or contracts. There are usually contact persons at universities and research institutions to support you with this as well.

Section Responsibilities and resources

Page Responsibilities

Question set Who is responsible for the adequate handling of the research data (description of the roles and responsibilities within the project)?

> **Help:** Possible roles are, e.g. project leader, data collector, data manager, data steward. You can find more examples on the [Harvard Medical School RDM Website](#)

Question: Who is responsible for the adequate handling of the research data (description of the roles and responsibilities within the project)?

Question set Who is responsible for curating the data after the project has ended?

Question: Who is responsible for curating the data after the project has ended?

Page Resources

Question: What resources are required to implement adequate handling of research data in the project?

> **Help:** This can involve technical or IT resources as well as expertise that is brought in, for example, by data management or IT experts. In case of additional costs for the use of infrastructure resources, please also enter these in the text field.