

The University of Bonn Research Data Service Center

Info: <https://www.forschungsdaten.uni-bonn.de>

Contact: forschungsdaten@uni-bonn.de

Guide to the 'Handling of Research Data' in DFG project proposals (Version 1.3 dated 21/09/2022)

This guide is intended as assistance with compiling the information on the “handling of research data” you are required to provide in DFG project proposals, under point 2.4. The guide is largely based on the [Research Data Handling Checklist](#) (version dated 12/21/2021) and the [Proposal Preparation Instructions](#) (last updated 09/2022). The guide draws on other DFG information as well, including the general [Guidelines for the Handling of Research Data](#) (dated 09/30/2015) and information on the [Handling of Research Data](#) provided on the DFG website (last updated 05/13/2022), including linked subpages. The guide is furthermore based on the [Guidelines for Safeguarding Good Research Practice](#) and the [FAIR principles](#).

Make sure to check whether any subject-specific requirements apply for your project! DFG review boards are increasingly releasing [subject-specific recommendations on the handling of research data](#). These documents are provided as further specification of or in supplement to the general DFG info. You should provide additional information accordingly if there are any such requirements applicable for your project.

The guide is tailored for offerings of the University of Bonn, pointing out numerous sources of further information as well as support services for research data management. For more information on data management plans and research data management for project applications, click [here](#). The [Research Data Service Center](#) will gladly assist you with determining the information you are to provide regarding the handling of research data. You can email us at researchdata@uni-bonn.de. For any questions regarding your project proposal which do not concern research data please contact University [Section 7.1, Research Funding](#).

In addition, we recommend that when dealing with research data, you consider the information and specifications provided by the relevant consortia of the National Research Data Infrastructure (NFDI). You can find an [overview of all NFDI consortia](#) on our website.

Model citation: The University of Bonn Research Data Service Center (2022): Guide to the 'Handling of Research Data' in DFG Project Proposals. V 1.3, online at: https://www.forschungsdaten.uni-bonn.de/de/files/leitfaden-zum-umgang-mit-forschungsdaten-in-dfg-projektantraegen/at_download/file

Usage license: CC0 (<https://creativecommons.org/publicdomain/zero/1.0/deed.de>)

1. Data description

How does your project generate new data? Is existing data reused?

Is data collected or generated internally within your research project, or is existing data reused? To promote the greatest efficiency of resources, a review should be made before the project begins as to whether existing research data can be reused for your project, eliminating the need for new generation. Further information on the gathering of research data is found [here](#).

When new research data is generated or collected: what methods or tools are used to generate the data?

If existing data is reused, who created the data? For what purposes and when? Where can the data be accessed?

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?

Specify the types of data utilized in your project. (e.g. numerical (databases, tables), text (documents), images, audio, video...). Please specify the file formats as well, if possible (often identified by the file extension, such as pdf, xls or txt). When selecting a data format, please take into consideration the potential consequences for collaborative use, archiving and reuse. It is therefore recommended that you select a standardized, non-proprietary format that is widely distributed within the specific community. This [list of recommended formats](#) provided by Data Archiving and Networked Services (DANS) may be helpful. (Please click [here](#) for further information about file formats)

Describe the planned work steps regarding the data—how is the data to be processed, analyzed and as applicable presented? Name/outline your methodological approach.

To what extent do these arise or what is the anticipated data volume?

State the volume of the datasets or estimate the accruing data volume (e.g. “less than 1 GB”, “0.5-1 TB” or “more than 100 TB”). Provide a breakdown by data type as appropriate.

2. Documentation and data quality

What approaches are being taken to describe the data in a comprehensible manner (such as the use of available metadata, documentation standards or ontologies)?

Please indicate where and how you will be conducting data documentation. The documentation should also enable third parties to audit and reproduce all work steps, including collection, processing, analysis, preparation, cleansing and presentation of the data. The meanings of all variables and codes in the dataset should also be clearly explained. Regular word processing or text editing software may be used to provide this information, for example in the form of a ‘readme’ file (file format txt). In some cases, special tools may be required (such as electronic lab journals). Please

click [here](#) for more information about data documentation. We provide a [readme template](#) for basic data documentation, which you are welcome to use for your datasets.

A well thought-out data organization strategy also contributes to the comprehensibility of the data, for example by defining clear file and folder naming conventions and directory structures, , or by introducing version control. More information on organizing research data can be found [here](#).

In many cases it is advisable to consider established standards, ontologies, controlled vocabularies and metadata schemas when structuring data and metadata. This increases the comparability and reusability of your research data. You should also utilize field-specific standards when possible, and consider the requirements for suitable repositories for taking over data after project conclusion (see section 5). Additional [notes on the assignment of metadata](#) and searching for [discipline-specific standards](#) are provided on our website.

What measures are implemented to ensure high data quality?

Outline for each dataset how data quality will be ensured, i.e. all measures concerning the content or formal controlling and maintenance of data and metadata, such as measures for:

- *Data completeness: is all information included? Are any files missing?*
- *Data consistency: are structure and formatting consistent throughout the dataset?*
- *Data integrity: is the data content correct and complete? What is the procedure in case of missing or incorrect data? Is there duplicate checking? Are data supplemented or enriched?*

Further information on quality assurance is provided [here](#) and [here](#).

Are quality controls in place and if so, how do they operate?

Specify who performs what measures when (data collection, compiling, processing, etc.).

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

Outline what is required in order to open and interpret your data. Is special software necessary, for example, or a specific system environment? This is particularly relevant if you use proprietary or non-standard data formats (see notes on data formats above).

3. Storage and technical backup during the project

How is the data to be stored and archived throughout the project duration?

Please indicate the systems in which your data will be stored during the project. The best locations for storing and editing data are primarily determinable based on your research data and workflows. Also, remember to make regular backups of your data. Cloud storage services are often recommended for automated synchronization of devices. Before using such services you should, however, carefully review the terms of use, server locations and other factors. The [Sciebo Sync & Share service](#) is recommended for University of Bonn staff and students. (For an overview of storage services available at the University of Bonn, please click [here](#).)

What is in place to secure sensitive data throughout the project duration (access and usage rights)?

Consider which persons require access to which data and at which times. In particular for projects with many participants, a targeted rights and roles management system is recommended for data access. Consideration may additionally have to be given to issues such as encryption and access controls for data requiring extra protection. Further information on data security is available [here](#).

4. Legal obligations and conditions

What are the legal specifics associated with the handling of research data in your project?

These may include data protection considerations in the event you will be collecting, processing or using personal data. Rules to be observed in processing personal data are set forth in the General Data Protection Regulation (GDPR) and the Data Protection Act of North Rhine-Westphalia. Personal data is defined as any information which refers or relates to an identified or identifiable living person. Such data may only be collected and processed on specific legal grounds, such as the informed consent of data subjects. Personal data require special protection against unauthorized accessing (e.g. through access control or encryption). When processed for research purposes personal data should be anonymized or pseudonymized as long as such is not in direct conflict with the research purpose. Further information on the use of personal data in research can be found [here](#).

In addition to data protection issues, patent law considerations and non-disclosure clauses may be of relevance, especially in projects with external partners.

Do you anticipate any implications or restrictions regarding subsequent publication or accessibility?

Generally, data should be shared as freely as possible—subject to legal and technological limitations. Accordingly, such limitations or prohibitions on sharing must be justified.

As a rule, personal data should be anonymized before sharing unless informed consent thereto has been obtained.

Third-party access to specialized data repositories may be regulated under a protection level scheme, protections ranging from required agreement to specific terms of use to permission for specific individuals to view data on-site on an encapsulated system. Embargo periods for the temporary retention of datasets may also be employed.

What is in place to consider aspects of use and copyright law as well as ownership issues?

In general, copyright is held by the person(s) who created the data. In Germany, the Copyright Act (UrhG) specifies that works of science that represent a “personal intellectual creation” are copyright

protected once a certain generative threshold is reached, meaning a certain creative activity can be recognized. Pure facts or data from measurement instruments do not qualify for copyright, although their compilation and editing do. If data is organized within a database, a property right similar to copyright (database right) may apply. Other legal areas such as ancillary copyright for press publishers and labor law may also be relevant. The employer (i.e. the University) generally holds research data usage rights, but in certain cases the author may hold such rights, such as with doctoral theses/dissertations and works produced by professors. The rule applies: The greater the extent to which a person is subject to direction in his/her research, the more usage rights will tend to be held by the employer. The employer always holds usage rights when data is generated in a situation of third-party research with external partners. Generally however, there are no problems associated with regular data sharing for research purposes in accordance with the FAIR data principles.

Data not subject to property rights is designated as “public domain.” Otherwise the copyright holder can negotiate a license for usage of the data with third parties. (further information on copyright and research data rights can be found [here](#))

Clarify in advance the legal terms of use for all data sets that you intend to reuse and/or generate. If reused data have an insufficient or no public usage license, you should obtain the permission of the copyright holder. If necessary, agree with your project partners on the terms under which the data should be passed on.

For information on legalities concerning research data and initial support, contact the Research Data Service Center. For formal guidance we refer you either to [Central Unit and Research Contracts](#) in Section 7 or to the [Data Protection Officer](#).

Are there any significant research codes or professional standards to be taken into account?

Check whether particular ethical aspects play a role for the research data in your project. For example, does your data have the potential to be misused for criminal or unethical purposes? Can the collection of data cause health, ecological, economic or other damage (see also [forschungsdaten.info: Ethical aspects in RDM](#))? Regarding ethical questions you may consult the [Research Ethics contact person](#) in Section 7.1, Research Funding.

5. Data exchange and long-term data accessibility

Which data sets are especially suitable for use in other contexts?

Please outline scenarios for potential reuse of your data and what user groups may be interested. In doing so, consider possible usage for other research purposes both in and outside your own discipline. Consider further usage possibilities. Could your data be of interest for teaching, for example, or do you have materials that could be beneficial in terms of methodology (survey methods, questionnaires, analysis techniques, visualization scripts, etc.)? Non-academic uses (for societal or business purposes) are also possible.

Which criteria are used to select research data to make it available for subsequent use by others?

Consciously select which data sets are suitable for subsequent use. In the case of larger amounts of data, a subset may be made available. The data sets selected should also be prepared so as to afford maximum clarity and transparency. In general, research data should be made accessible via a repository. Indicate the repositories suitable for receiving your data and what requirements apply. (further information on data publication is found [here](#). For information on locating discipline-specific repositories consult these [discipline-specific information pages](#) or the relevant [NFDI consortia](#). . University of Bonn staff and students can use the [Repository RADAR](#) for this purpose (remember to file for reimbursement of any accruing costs; see section 6).

You should furthermore agree in advance with all project participants involved on the terms of use for third parties and on the license for your data. . The [DFG recommends licenses structured to be as open as possible](#). [Creative Commons](#) and [Open Data Commons](#) licenses are suitable for many datasets. Data that is not subject to copyright should be published in the public domain. Please note that there are dedicated open source licenses for software, such as the [GNU General Public License](#) and the [MIT License](#). Please click [here](#) for further information about licenses.

Are you planning to archive your data in a suitable infrastructure? If so, how and where?

Your answer to this question should normally be “yes”. In line with the [Guidelines for Safeguarding Good Research Practice](#), [the DFG expects](#) primary data forming the basis for a publication to be stored long-term (i.e. for at least ten years) “in the researcher's own institution or an appropriate nationwide infrastructure”. The aim is to make the research results comprehensible and verifiable.

Some repositories enable restricting data access so that the data are searchable, i.e. their existence verifiable, but are only accessible on request, with appropriate security measures in place. Such non-public archiving is possible for members of the University of Bonn via the [RADAR repository](#). If the integrity and accessibility of the data is guaranteed, it can also be archived for 10 years in a [storage service](#).

Are there any retention periods?

Many repositories offer the option of an embargo deadline whereby the date of publication can be delayed as necessary.

When is the research data available for use by third parties?

Please indicate the planned date of publication for all data sets that you will be publishing. The data should always be published as quickly as possible. Data that serves as the basis for single publications should ideally be published at parallel to that work.

6. Responsibilities and resources

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?

What expenses will be incurred for research data management for your project, justifying the corresponding project funding? Examples: expenses for setup and maintenance of special IT infrastructure for your project, dataset preparation, data repository user fees. Such expenses are generally eligible to be covered under the project funding applied for. During the planning phase, create a list of costs to ensure the feasibility of your project. For example, the [DFG provides funding](#) for “staff costs, project-related hardware and software, and usage fees”.

Who is responsible for curating the data once the project has ended?

As applicable, name the individuals responsible for ongoing data maintenance and processing beyond the project term. External data archives may provide such services in some cases.

If no further measures are taken regarding the data after ending of the project, name contact persons from whom information on the datasets can be obtained.

Additional note: supporting institutions

Under the [DFG proposal preparation instructions](#), you are supposed to outline “how the institutions involved in the project will contribute to data and information management”. In doing so you may refer to the [Research Data Service Center as an advisory unit](#), after consultation with us. Other basic services, like storage options or a research data repository, are listed on our website. Depending on the topic and specialist orientation, other institutions of the University of Bonn may also be able to provide assistance, such as the [Bonn University and State Library](#), [University IT](#) or the [Bonn Center for Digital Humanities](#). Regarding any legal questions you may contact us at [Central Unit and Research Contracts](#) in Section 7, or the University’s [Data Protection Officer](#) depending on the nature of the matter. Regarding ethical questions you may consult the [Research Ethics contact person](#) in Section 7.1, Research Funding. External contact persons, such as the data centers and specialist information centers, can support with RDM.