# Data management

General information:

* This checklist is based on recommendations on the handling of research data for researchers wishing to submit grant applications, including information on how to plan and describe the handling of research data as part of research projects, from the German Research Foundation (DFG): <https://www.dfg.de/download/pdf/foerderung/grundlagen_dfg_foerderung/forschungsdaten/forschungsdaten_checkliste_en.pdf>.
* Please also note the subject-specific recommendations: <https://www.dfg.de/en/research_funding/principles_dfg_funding/research_data/recommendations/index.html>.
* This checklist contains ideas and different aspects relating to data management that should be taken into consideration before starting and while working on a project.
* It is possible that not all of these aspects apply to your data.
* It is not necessary that you have fully detailed solutions for all aspects when applying. However, you should have given some thought to relevant issues and make corresponding statements in your application.
* The checklist is an aid that will help you put a focus on data that is reused or generated in the context of the project and consider this data from different perspectives. The objective is to document the research process in a way that provides a basis for the data to be subsequently reused.
* Depending on the scope of the project, the completed checklist will be between one and two pages in length (excluding the blocks of tables).
* **Tip:** By deleting the individual cells of the tables and removing the page breaks, you can produce a completed checklist in continuous text form that you can attach as a separate document or copy and paste into your application.
* **Note:** The glossary from forschungsdaten.info may be helpful for the description of individual terms: <https://www.forschungsdaten.info/glossary/>.
* **If you have any questions, please do not hesitate to contact us via:**

[rds@uni-due.de](mailto:rds@uni-due.de) | <https://www.uni-due.de/rds/en/index.php>

|  |
| --- |
| **Original questions on data description from the DFG checklist**  How does your project generate new data?  Is 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? |
| **Notes and examples from UDE relating to data description**  Describe the data that will be used and newly generated in the context of your project here. State whether you are going to reuse existing data or generate data yourself and briefly describe the relevant data. In doing so, specify the types/formats of the data and their subsequent use, as well as the anticipated overall volume of the data. Also indicate which methods and tools are required to use the data [the question regarding methods and tools (e.g. software) can also be answered in the second section].  ***Notes:***  *Provide a summary of the data generated in the course of the project (e.g. by data type such as survey data or measurement data, or by special characteristics of the data such as sensitivity) and describe it briefly with regard to formats, further use within the project, (overall) volume and the software used (for creation and subsequent use).*  *Is special software required?*  *If you reuse existing data, provide information on the source and the terms of use (e.g. Creative Commons licences) in addition to the description.*  ***Examples:***  *Based on recent searches in subject-specific data archives (repositories), data that is relevant to the issues addressed in the project is not available. Therefore, the following data collection processes are to be conducted: ...*  *The data collected will be stored along with the relevant metadata in the form of a CSV file. The data will be analysed using the XY statistical software. The volume of the data collected in the course of the project is estimated to amount to a total of ... GB.*  *The survey data can be accessed using standard spreadsheet software.* |

## Data description

Room for your text.

|  |
| --- |
| **Original questions on documentation and data quality from the 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 adopted 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? |
| **Notes and examples from UDE relating to documentation and data quality**  Explain your approach to describing the data generated or reused within the project in a way that makes their collection and use transparent here.  In addition to the documentation process, also provide information on how the quality of the data is ensured and controlled. In doing so, also indicate which methods and tools are required to use the data [the question regarding methods and tools (e.g. software) can also be answered in the first section].  ***Notes:***  *Do you document how data is generated and used, for example in the form of readme files or electronic (laboratory) notebooks (see also:* [*https://wissenschaftliche-integritaet.de/en/code-of-conduct/documentation/*](https://wissenschaftliche-integritaet.de/en/code-of-conduct/documentation/)*)?*  *Do you use existing metadata standards in your descriptions (e.g. as a result of repository requirements)?*  *Is there any sensitive data that must be documented in anonymised form? If that is the case, how is this done?*  *What quality assurance mechanisms are applied (see also:* [*https://wissenschaftliche-integritaet.de/en/code-of-conduct/cross-phase-quality-assurance/*](https://wissenschaftliche-integritaet.de/en/code-of-conduct/cross-phase-quality-assurance/)*)?*  ***Examples:***  *The standards for storing, naming and versioning data, as well as the metadata to be compiled, will be agreed with the project partners at the start of the project. The objective is for the metadata to be compiled in such a way that it describes the data suitably for the relevant discipline and that it is machine readable and interoperable.*  *We aim to apply the following measures to ensure the quality of the data: ...*  *All steps of data processing will be documented [e.g. in an electronic lab notebook] as quickly as possible.*  *The transcription and annotation of audio and video clips will be double checked.* |

## Documentation and data quality

Room for your text.

|  |
| --- |
| **Original questions on storage and technical archiving throughout the project from the DFG checklist**  How is the data to be stored and archived throughout the project duration?  What is in place to secure sensitive data throughout the project duration (access and usage rights)? |
| **Notes and examples from UDE relating to storage and technical archiving throughout the project**  Describe here where the data will be stored and who will have access to it in the course of the project. You may want to differentiate this information by data type. If sensitive data is generated during the project, please provide specific information as to how access and user rights are managed for this data.  ***Notes***:  *See UDE’s storage matrix:* [*https://www.uni-due.de/rds/en/daten\_speichern.php*](https://www.uni-due.de/rds/en/daten_speichern.php) *or* [*https://www.uni-due.de/rds/en/speicher\_matrix.php*](https://www.uni-due.de/rds/en/speicher_matrix.php)  *Are there any special circumstances, e.g. large volumes of data (> 50 GB), that would need to be handled separately?*  *Examples:*  *During the course of the project, the data will be stored by the University of XY’s data centre.*  *Sensitive data excepted, data will be exchanged via a Sciebo project box throughout the project.*  *With the exception of ... data, all researchers within the project will have access to all data. Only the following individuals will have access to the ... data: \*given name\* \*surname\*; \*given name\* \*surname\*.*  *For the project duration, it is expected that a data volume of ... GB will need to be stored.* |

## Storage and technical archiving throughout the project

Room for your text.

|  |
| --- |
| **Original questions on legal obligations and conditions from the 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? |
| **Notes and examples from UDE relating to legal obligations and conditions**  Describe the legal specificities of the data here. In particular, elaborate on potential opportunities for subsequent publication and/or accessibility. In doing so, also consider aspects of copyright and data protection. Are there any important research codes or professional standards that should be taken into account?  ***Notes:***  *See also:* [*https://wissenschaftliche-integritaet.de/en/code-of-conduct/legal-and-ethical-frameworks-usage-rights/*](https://wissenschaftliche-integritaet.de/en/code-of-conduct/legal-and-ethical-frameworks-usage-rights/)  *Further information on codes and professional standards:* [*https://wissenschaftliche-integritaet.de/en/comments/further-links-relating-to-legal-and-ethical-frameworks/*](https://wissenschaftliche-integritaet.de/en/comments/further-links-relating-to-legal-and-ethical-frameworks/)  *Is there any sensitive data?*  *Do any data protection requirements have to be met? See also:* [*https://www.uni-due.de/verwaltung/datenschutz/*](https://www.uni-due.de/verwaltung/datenschutz/) *(German only)*  *Does an ethics committee need to approve the collection of data?* [*https://www.uni-due.de/ethikkommission/*](https://www.uni-due.de/ethikkommission/) *(German only)*  *Are there any copyrights or patent rights that need to be considered?*  *Who owns the rights of use of the data (the researchers? the projects? the employer?)?*  *In case of cooperation with businesses: Are any agreements in place regarding the use and potential publication of the data?*  ***Example:***  *The ... data is protected by copyright.*  *The University of Duisburg-Essen has been granted a non-exclusive right to use the ... data.*  *There are no data protection-related issues with regard to conducting the research project.*  *The project has been examined by an ethics committee and ethical clearance has been granted.*  *All parties involved in the project will sign a cooperation agreement at the start of the project in which they grant one another non-exclusive rights of use of all research data to be generated during the project. This includes the right to publish (pseudonymised) data.*  *Approval from the authorities is not required for the planned study.* |

## Legal obligations and conditions

Room for your text.

|  |
| --- |
| **Original questions on data exchange and long-term data accessibility from the 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? |
| **Notes and examples from UDE relating to data exchange and long-term data accessibility**  Explain here whether you are planning to publish data during the course of the project or after the project has been completed and which criteria will be applied in selecting the relevant data. Please also describe whether and where the data is to be archived and whether any retention periods are in place that have to be respected if the data is to be accessible to third parties.  ***Notes:***  *See also:* <https://wissenschaftliche-integritaet.de/en/code-of-conduct/providing-public-access-to-research-results/>  *See also:* [*https://wissenschaftliche-integritaet.de/en/code-of-conduct/archiving/*](https://wissenschaftliche-integritaet.de/en/code-of-conduct/archiving/)  *Are the selected file formats suitable for long-term archiving?*  *Based on which criteria are datasets selected for publication?*  *On data publication (see also):* [*https://www.uni-due.de/rds/en/repositor.php*](https://www.uni-due.de/rds/en/repositor.php)  ***Examples:***  *The dataset used to investigate XY could be relevant for other research institutions. Therefore, the data and metadata are to be made available for subsequent use in a subject-specific repository.*  *The following project data cannot be published and made available for subsequent use for reasons of competition:*  *The findings of the project are to be published in at least two open access publications. All project data that is relevant to these publications, including the corresponding metadata, is to be archived and kept available in a subject-specific repository (e.g. the XY repository) for a period of ten years.*  *The project data will be assigned an identifier (DOI) and made available in the XY University’s institutional repository in a freely accessible manner when the project findings are published or, at the latest, when the final report is submitted six months after the end of the project.* |

## Data exchange and long-term data accessibility

Room for your text.

|  |
| --- |
| **Original questions on responsibilities and resources from the 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? |
| **Notes and examples from UDE relating to responsibilities and resources**  Explain here who is to be assigned which responsibilities relating to the data and in what role during the project and after the end of the project. Also provide a rough estimate of the resources (time, costs, etc.) this will require, also taking into account the costs that may be incurred for preparing and archiving the data.  ***Notes:***  *See also:* [*https://wissenschaftliche-integritaet.de/en/code-of-conduct/stakeholders-responsibilities-and-roles/*](https://wissenschaftliche-integritaet.de/en/code-of-conduct/stakeholders-responsibilities-and-roles/)  *Who is responsible for the generation and documentation of data?*  *Who is responsible for preparing the data for publication/provision/subsequent use?*  *Who is the contact responsible for the data after the end of the project?*  ***Examples:***  *The expenses/costs for archiving the data over the course of the project are anticipated to amount to €... . The data volume to be archived after the end of the project is estimated to amount to ... GB. The research unit will cover the expenses/costs for this.* |

## Responsibilities and resources

Room for your text.