Example Data Management Plan: DFG Checklist for Handling Research Data.

1. Data description

The planed research data within the [XYZ project] will be collected via an online questionnaire. The software [X.2] provided by the Computer and Media Service (CMS) of Humboldt-Universität zu Berlin is used for this purpose. The survey data is analyzed using the open source statistical software R and is saved in the form of the data set (csv), the R analysis script (r) and a series of graphics (tiff). Furthermore, a readme file (txt) as well as the questionnaire (pdf/a) and a codebook (pdf/a) are created to describe the data. In addition to the self-collected research data, publicly available data in the form of public statistics (csv), reports (docx, pdf) and legal regulations (html, pdf) will be reused and referenced for the project. The expected total size of all files is a maximum of 50 GB.

2. Documentation and data quality

Metadata is created via the web form of the [X institute] in accordance with the Domain-Specific Standards. Additional documentation of the research data in the form of a readme, questionnaire, codebook and R syntax is also planned. Keywords are assigned according to the subject-specific thesaurus. The data will be classified in the web form using the subject-specific classification of [domain/field]. Quality controls are embedded within our data collection, processing, and analysis workflows, with checks and validations at each stage to identify and rectify any errors or inconsistencies. This includes automated validation scripts and manual review processes. We use the [XYZ] software packages, analytical tools, and programming environments for data processing, visualization, and statistical analysis.

3. Storage and technical archiving the project

Secure storage and backup will be ensured during the project period by the project management in co-operation with the responsible data processing officer of [Institute Y]. The university's cloud-based storage "HU-Box" is used for the storage and collaborative processing of data during the project. This enables clear access management and simple user administration. Access to sensitive data is restricted to authorized personnel only. Encrypted, password-protected folders are used for sensitive data, which can only be viewed and processed by authorized employees. An automated backup is performed every night.

4. Legal obligations and conditions

[We anticipate [no] potential restrictions regarding subsequent publication or accessibility of our data.] As part of the online questionnaire, participants will be informed about subsequent publication while maintaining anonymity. The online survey is created in compliance with the GDPR with the advice of the institutional data protection officers. This includes informed consent from the respondents and separate consent for the subsequent publication of the data collected. An ethics vote is obtained beforehand from the responsible ethics committee at Humboldt-Universität. A cooperation agreement will be concluded with [project partner Z] to clarify the copyright of the data and a data management plan will be drawn up for the project.

5. Data exchange and long-term data accessibility

In addition to being analyzed directly by the project group, the dataset will also be relevant for other research projects. As no comparable data is available for secondary analysis, the research data, the R analysis scripts and the associated questionnaire will be made available to other researchers at the [X institute] under a CC-BY licence. The study receives a Digital Object Identifier (DOI) via the [X institute] data archive. As required for good scientific practice, the project results and all relevant research data will also be stored at the Humboldt-Universität zu Berlin for at least ten years. For this purpose, a procedure for transferring the data to the CMS backup service will be agreed with the institutional data processing officer. The [X institute] for the Social Sciences will curate the data after the end of the project.

6. Responsibilities and resources

In accordance with the research data policy at Humboldt-Universität zu Berlin, the project manager is responsible for all aspects of research data management. However, individual sub-areas may be delegated to the project staff. For example, the preparation of research data for publication in the repository is planned for [project member #3]. The [X institute] will make the data available and archive it free of charge after consultation with the repository. The long-term storage of the data is free of charge by the CMS of the HU Berlin.