Data Policy RTG (Example 1)

ID

Funder

Deutsche Forschungsgemeinschaft (DFG)

Project Name

Research Training Group

Project Description

The aim of the RTG is to create a research environment for PhD students and to provide structured training for young scientists. The research projects focus on …

The studies within the RTG include … [data types and research types]

Principle Investigator

Speaker of the RTG:
PIs:

PhD-Students

Employed and associated PhD Students of the RTG

Project Data Contact

[RTG Coordination Office]

Term of the Project

Date of First Version

Date of Last Update

Related Policies

Leitlinien zum Umgang mit Forschungsdaten (DFG)
Verwendungsrichtlinien Graduiertenkollegs
Research data policy of the Georg-August University Goettingen

Definition

"Data" shall mean any information in analogue or digital format including writings, drawings, descriptions, recordings and software, which is produced, used, acquired, derived from data or stored in the course of conducting research within the RTG research projects by any of its participants.
"Meta-data" shall mean any data describing and documenting such data.
"Data supplier" is defined as the researcher who conducted the study (generally the PhD student).
"Raw data" is defined as the data generated during data acquisition. Raw data can already be anonymized for privacy protection.
"Project" means a completed research study.
"RTG Members" include the listed PIs and the PhD students (employed or associated).
"Long-term storage" means the storing of data in an archive at the GWDG or the DPZ IT Service for at least 10 years after publication.
"Storage" is defined as any data saving procedure.

Period of data retention

All data in relation to publications is to be retained for at least 10 years after publication.

Expected data

The expected data types will cover:
Audio samples (Database samples and own recording)
Photos/picture (Database samples and self-created material)
Questionnaire data
…

The expected data amounts are:
[daily/monthly/yearly/total]

The data are accumulated during the period of the RTG.

For the data collection with human participants the ethic regulations of the DGPs do apply. Human participants were informed about the data recording procedure and the storage regulation. They gave written informed consent prior to the data collection. The data supplier and the PIs are responsible for complying with the ethics regulations. Ethic assessments, to secure that the projects comply with the ethic regulations, can be obtained from the ethics committees of the institute of psychology (University of Göttingen) or of the University Medical Center. Existing databases are used as stimulus material. These databases are either open source for academic purposes or written permission has been given to use the data base as stimulus material.

Data storage and Data formats

Digital data are stored using the facilities of the GWDG, the IT provider of the University of Göttingen, or of the project's home institute. Non-digital data, such as questionnaires, lab books and informed consent sheets, are stored at the respective departments. The IT providers are responsible for the data integrity,

security and for providing enough storage space to store the data. The research departments, in which the research is conducted, are responsible for arranging the storage space. All costs which incur in connection with the storage shall be borne by the respective research departments. PhD students and their supervisors are responsible for storing the data in a way that they are complete (including all information to reproduce the project), retrievable, accessible and comprehensible in a way that the data can be provided to others without further explanations. The RTG Coordinator will give trainings and assistance for the storing.

For digital long-term storing, researchers are asked to convert the data to open formats optimal for preserving. The following formats are recommended:
Audio files - .wav
Video - .mpg4
Photos, figure, picture - .tif
Numerical data - .csv
text - .txt
Wherever proprietary formats are necessary, information on the appropriate software and the supplier should be provided. If in accordance with copyright regulations, a copy of the software should be deposited.

English should be used for any labels and information. Folders and files should have self-speaking titles, including the date of last version.

Meta-data

Every project must include extensive meta-data with all details required to understand the study (project-related meta-data) and the single data files (file-related meta-data). The Dublin Core Metadata Element Set is used as the meta-data standard. The coordinator provides a .csv template sheet on a joint SharePoint. The meta-data include the description, where to find all data related to one project and is stored together with the data. The project related meta-data is additionally stored centrally at the SharePoint. All members of the RTG have read permission to the meta-data entries on the SharePoint. The coordinator is responsible for training the PhD Students and for checking the entries. Meta-data entry on these sheets is done constantly during all steps of data processing and publication.

Organization of Data Storage

Data deposition on a long-term storage for every project comprises the raw data (fixed datasets), derived data (analyses and processing), and meta-data. In case a publication or project is based on existing data sets that are long-term stored already, a reference to the raw data set in the meta-data would be sufficient.

Data deposition is to be arranged as follows:

1) During data collection and analysis, data can be saved on the personal storage used at the department, which is provided, secured and backed up by the IT provider (GWDG account at the university, personal storage on the server at the project's institute). Data collected outside of the institution should be backed up using two different hard drives. After the return, data must be directly secured on the personal institutional file server. Data must not be stored using Dropbox, or other commercial or third-party cloud storage services, due to the lack of data security. The GWDG ownCloud service is recommended as alternative cloud storage solution. Meta-data entry using the meta-data sheets should be done in parallel to data collection and analysis. During the research phase, PhD students and their PIs are responsible for the data storage.

2) All digital data regarding a finished project must be deposited to a long-term storage, either at the IT Service provider of the University or the project's institute, or open data repositories. Non-digital documents, such as pen and paper questionnaires, lab books, or informed consent sheets must be long-term stored at the department, in which the study was conducted. The long-term storage and the meta-data entry on the SharePoint, incl. how to access the long-term stored data, has to be completed when handing in the dissertation, or with finalization of other research output (paper) or with employment contract expiry in other cases, whatever occurs first. The PhD student and the PIs are responsible for submitting the data to a long-term storage. The service provider (either the GWDG, the institute's IT Service or the open data repositories), is responsible for the long-term storage.

3) Compliance of the centrally located meta-data is controlled by the data supplier and is regularly (at least twice a year) carefully reviewed by the RTG Coordinator. If he/she considers that documentation is not sufficient, the data supplier will be requested to complete meta-data entries.

Protection

Personal data must be anonymized, thus no reference is existent between stored data and participants. The anonymization has to be done as soon as possible, if applicable directly after data collection, but latest for the data storage. Videos and voice samples that cannot be anonymized should be long-term stored in an encrypted folder.

Data Dissemination

1) The outcome of the projects will be made publicly available through publication in journals, as book chapters or as comparable works. Researchers are strongly encouraged to publish in open access journals; otherwise publications must be uploaded to open journal repositories as soon as possible.

2) Data can be exchanged within the RTG with agreement of the data supplier (in case he/she has not left the institute) and the PI of the project.

3) Data will be made available to people outside of the RTG on email request, if not stored openly in data repositories, and data use must be based on a written agreement between the PI of the project, the data user and the data supplier (in case he/she has not yet left the institute), in which the use of the data is regulated. If data is published in data repositories the regulations of data use of the repository applies.

4) Accessed data must be used for scientific purposes only, i.e., commercial use of data is not allowed.

5) The usage of the data must be appropriately acknowledged and agreements on co-authorship of publications must follow the rules of good scientific practice. Before anything is submitted for publication, the PIs of the involved projects must be consulted.

6) Publications must acknowledge that the data were compiled in the RTG funded by the German Research Foundation (DFG) with the standard sentence: This research was supported by the German Research Foundation (DFG) as part of the RTG.

7) This protocol does not affect any rights of RTG members under applicable copyright legislation. As far as project data are subject to third party's copyright, such right has to be respected.