



University of Potsdam

**SFB
1287**

Limits of Variability in Language
Cognitive, Grammatical, and Social Aspects

Data Management Plan

Editorial

<i>Project Number:</i>
<i>Name of Experiment / Acronym / Number:</i>
<i>PI or responsible person (head of the study):</i>
<i>Filled out by:</i>
<i>Last update:</i>
<i>Version:</i>
<i>Storage location of this DMP:</i>
<i>Update reason:</i>

Table of Contents

Editorial..... 2

Table of Contents..... 3

General Information 4

Technical Information..... 10

Publication 15

Legal and Ethics 19

Storage and long-term preservation 23

References..... 25

General Information

1. Are there requirements regarding the data management from your scholarly/scientific community?

2. If yes, what are the requirements?

General Information

3. What kind of dataset is it?

The following questions collect information on the data that is produced or used in the project. They also help to estimate the value of the data in terms of potential re-use and long-term preservation. In the case of personal data, the principle of data minimization (Art. 5 EU General Data Protection Regulation) allows the collection of personal data only when there are no other reasonable means to clarify the research question (re-use of existing data would be such a reasonable means). Also, there shall be no more information collected than necessary. The information regarding the data collected, produced or used in the project is gathered along the datasets. The definition of these datasets is an important conceptual decision that has to be made individually and carefully for each project.

Please fill in the form for each dataset separately:

Name of Experiment / Acronym / Number:

PI or responsible person (head of the study):

Other persons involved:

General Information

Subject area:

Method / Type of data:

Anonymizable data?

Participants:

General Information

Short description (of the study):

Other comments:

Time for data collection (approximate)

Time for data analysis (approximate)

General Information

Related publications:

Keywords (used in publications):

Link to preregistration:

Funding reference:

General Information

4. Which individuals, groups or institutions could be interested in re-using this dataset? What consequences does the reuse potential have for the provision of the data later?

It is important to specify whether the data will be permitted for reuse. But legal impediments, such as privacy, and copyright must be taken into account.

Technical Information

1. Where is the dataset stored during the project?

Note:

Please delete all project-files from source that is not part of the University of Potsdam.

for raw data

SFB1287 - File Server (backed up)

SFB1287 - "Analysis Database" (backed up)

SFB1287 - "Working Memory Database" (backed up)

Box.UP - Cloud (University of Potsdam) (backed up)

Git.UP - Cloud (University of Potsdam) (backed up)

Researcher's Computer

Computer in the laboratory

External Hard Disk Drive

Ibex Farm Website

Personal Google Cloud

Personal GitHub

Personal GitLab

Personal Dropbox

Personal OwnCloud

Technical Information

for analysis data

SFB1287 - File Server (backed up)

SFB1287 - "Analysis Database" (backed up)

SFB1287 - "Working Memory Database" (backed up)

Box.UP - Cloud (University of Potsdam) (backed up)

Git.UP - Cloud (University of Potsdam) (backed up)

Researcher's Computer

Computer in the laboratory

External Hard Disk Drive

Ibex Farm Website

Personal Google Cloud

Personal GitHub

Personal GitLab

Personal Dropbox

Personal OwnCloud

Technical Information

for further documentation, related code or software

SFB1287 - File Server (backed up)

SFB1287 - "Analysis Database" (backed up)

SFB1287 - "Working Memory Database" (backed up)

Box.UP - Cloud (University of Potsdam) (backed up)

Git.UP - Cloud (University of Potsdam) (backed up)

Researcher's Computer

Computer in the laboratory

External Hard Disk Drive

Ibex Farm Website

Personal Google Cloud

Personal GitHub

Personal GitLab

Personal Dropbox

Personal OwnCloud

Technical Information

2. If data is stored on lab or personal computers, please describe the backup strategy.

Technical Information

3. Which file formats are used?

When choosing a data format, one should consider the consequences for collaborative use, long-term preservation as well as reuse. It is advisable to use formats that are standardised, open, non-proprietary and well-established in the respective scholarly community. A table with recommended file formats can be found in found in [Kristin Briney, Data Management for Researchers, Pelargic, 2015, pages 133-134.](#)

Publication

1. Will this dataset be published or shared?

anonymizable data

- *ID will be removed (and or code-list will be destroyed) [legally correct: code list will be destroyed as soon as possible without jeopardizing experiment; exception: follow-up study planned, if so, definitely talk to UP data protection officer (Dr. David Kneis; datenschutz@uni-potsdam.de) on how to do this correctly]*
- *Publication of anonymized data and code on OSF or RADAR (University of Potsdam) (or as required by the Journal)*

non-anonymizable data

- *on RADAR (University of Potsdam) but not accessible from the outside world*

Publication

2. If yes, under which terms of use or license will the dataset be published or shared?

The options refer to the licenses of the [Creative Commons](#) family. If data is anonymised / pseudonymized, it's probably not legally required, but might be good in terms of research ethics to adjust consent forms / subject information sheets.

for data

for scripts

for software

Publication

3. If yes, when will the data be published?

Recommended procedure: Upload data and obtain digital identifier (e.g., DOI, OSF link) when submitting the first paper; thus, you can cite the data in the paper. If necessary, restrict public access (embargo) until last paper published (max. 2 years).

4. If no, please explain why not. Please differentiate between legal and contractual reasons and voluntary restrictions.

Publication

5. Which measures of quality assurance are taken for this dataset?

Legal and Ethics

1. Does this dataset contain personal data?

The EU General Data Protection Regulation (GDPR) defines in Art. 4 personal data as "any information relating to an identified or identifiable natural person". An identifiable natural person is "one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person".

For advice and/or trainings on how to comply with privacy regulations - including proper anonymization & pseudonymization - you can always contact UP's privacy officer, Dr David Kneis, at datenschutz@uni-potsdam.de. As the privacy officer, his perspective is focussed more on the legal side of things than the research ethics or technical aspects.

Legal and Ethics

2. Are the data anonymised?

***Anonymised data:** ID will be removed (and or code-list will be destroyed) [legally correct: code list will be destroyed as soon as possible without jeopardizing experiment; exception: follow-up study planned, if so, definitely talk to UP data protection officer on how to do this correctly]*

for raw data

for analysis data

for published data

Legal and Ethics

3. Does the project use and/or produce data that is protected by intellectual or industrial property rights?

Measurement data has no intellectual property, so usually, the answer here will be „no“. Data or software can be subject to intellectual or industrial property rights. Applicable laws differ broadly even within EU. According to the German copyright law (UrhG) works of literature, scholarship and the arts that can be regarded as a “personal intellectual creation” are protected by copyright. Mere data, e.g. measured data or survey data, and metadata (except in some cases descriptive metadata) are not protected by copyright. § 2 of the UrhG lists the following kinds of protected works (list is not concluded):

- *linguistic works such as written works, speeches and computer programs*
- *works or the fine arts including works of the applied arts as well as sketches of such works*
- *works of photography*
- *descriptions and illustrations of scholarly or technical nature such as drawings, plans, maps, sketches, tables and three-dimensional representations*

According to § 3, copyright is also applicable to translations and other modifications or adaptations of work if they are individual intellectual creations of the editor. Finally, according to § 4 copyright also extends to collected editions and database works. Collected editions are: “collections of work, data or other independent elements that are individual intellectual creations based on the selection and arrangement of the elements”.

Database works are defined as “collected editions, the elements of which are arranged in a systematic or methodical way and can be accessed individually by electronic means or in other ways”.

Legal and Ethics

4. If yes, please explain which!

Storage and long-term preservation

1. Does this dataset have to be preserved for long-term?

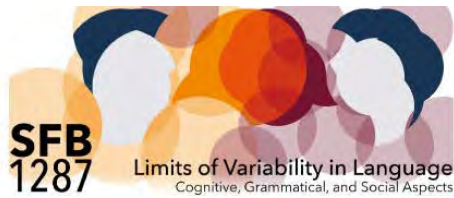
The DFG expects primary data that is the basis of a publication to be stored in the researcher's own institution or an appropriate nationwide infrastructure long-term (for at least 10 years).

2. What are the reasons this dataset has to be preserved for the long-term?

Storage and long-term preservation

3. Where will the data (including metadata, documentation and relevant code) be stored or archived after the end of the project?

References



SFB 1287

<https://www.uni-potsdam.de/de/sfb1287>



University of Potsdam

<https://www.uni-potsdam.de>



Deutsche Forschungsgemeinschaft e.V.

<https://www.dfg.de>