

Archiving publication packages at the Institutes of Psychology and Education & Child Studies¹

What is a publication package?

A publication package is an anonymized/pseudonymized bundle of data, that is created with the [FAIR principles](#) in mind. This package contains all the materials necessary to replicate the study, reproduce the results and reuse the data. For all published empirical studies and empirical studies appearing in (unpublished) PhD thesis chapters and research master theses, a publication package has to be stored on the platform DataverseNL for at least ten years.

What should a publication package contain?

- 1. Manuscript or publication:** for publications either the Original Author Manuscript as submitted to the publisher (preprint) or the Author's Accepted Manuscript (postprint). You may use the published manuscript only if the publication was originally open access on the publisher's website.
- 2. The instructions, procedures, the design of the experiment and stimulus materials** (interview guide, questionnaires, surveys, tests) that are necessary for an exact replication of the research.
- 3. The (de-identified) raw data files, documentation of steps for de-identification, and blank consent form.** If the raw data files have been accessibly stored in an external archive (e.g., at DANS), making reference to the files in this archive will suffice. Such externally archived raw data may include primary or secondary data. Raw data files provide the most direct registration of the behaviour or reactions of participants (e.g., unfiltered export file of online surveys, raw time series for an EEG measurement, e-dat files for an E-Prime behaviour experiment, recordings or transcripts of interviews, descriptions of observations, archive and other source or media material).
- 4. Computer code** (e.g., SPSS syntax file, MATLAB analysis scripts, R code) describing all steps taken to process and convert the raw data into the published results, including explanations of the steps in English.
- 5. The processed data files** that were eventually analysed when preparing the article (e.g. an SPSS data file after transforming variables, after applying selections, etc.).
- 6. The data management plan** if this was created for the study.
- 7. A readme file** (metadata) describing how data files are organized and internally structured, as well as any other information that a researcher in a similar field would need to reproduce the published results or carry out new analyses based on the components of the publication package. The readme file must also contain:
 - a. Name of the person who stored the documents or files;
 - b. Division of roles among authors, indicating at least who analysed the data;
 - c. Date on which the manuscript was accepted, including reference;
 - d. Date/period of data collection;
 - e. Names of people who collected the data;
 - f. If applicable: addresses of field locations where data were collected and contact persons (if any)
 - g. Whether or not an ethical assessment took place before the research, and, if relevant, study reference from and statements made by the Ethics Review Committee;
 - h. Whether the data is made open, restricted, or closed. A valid reason for not opening up the data needs to be included;
 - i. A list of excluded subjects, corrupted data files, and any other exceptions to the standard analysis pipeline, to enhance understanding of the data and the data analysis.
- 8. Documents related to the ethical approval** or a reference to such documents.

¹ This instruction is based on the Guideline for the archiving of academic research for Faculties of Behavioural and Social Sciences in The Netherlands. Version March 2022.

How to publish your publication package?

A publication package must be stored within one month after the definitive publication of the manuscript. The data stewards are there to support the researchers in publishing their publication packages on DataverseNL. In order to publish a publication package on DataverseNL, researchers are asked to:

1. Prepare the components listed above in the following way: Each of the 8 components should consist of one file or one zip file (including multiple files) except in cases where there are multiple types of data files (e.g., E-Prime, EEG, heart rate); in that case make one zip file per data type. Number the 8 components so they appear in the correct order in DataverseNL.
2. Encrypt the publication package with a unique password and then place it in J:\ResearchData\FSW\Archivering DataverseNL\ (choose folder of unit). Encryption can be done with the program 7-ZIP (can be installed via Software Center). Right-click the publication package folder, via 7-Zip, select 'Add to archive...'. Select '.zip' as 'Archive format', and enter a password.
3. Fill in this [Formdesk questionnaire](#) with relevant metadata about the publication package.
4. Inform the Behavioural Sciences data stewards that a publication package is ready to be checked via datastewards_psy_ped@fsw.leidenuniv.nl, don't forget to add the password too.

The data stewards will check if the publication package consists of all 8 components and does not contain personal data. The data stewards will notify the researcher if the package does not meet these aspects. If the package does meet these aspects, the data stewards will upload the package in DataverseNL, after which the researcher receives the DOI-link.

'As open as possible, as closed as necessary'

In line with the Netherlands Code of Conduct for Research Integrity, the default in DataverseNL is that all components of the publication package are 'open' (i.e., openly available for everybody with a DataverseNL account). If there is a valid reason, the researcher could choose between 'restricted' (not directly available, but request for access can be sent to the contact person via mail) or 'closed' (never available for anyone, and is only added for archiving purposes). A valid reason for choosing the 'restricted' or 'closed' options should be provided in the readme file and communicated to the data stewards.

New versions of an earlier uploaded dataset

When a publication package has been published on DataverseNL, it can't be changed anymore. If errors or mistakes are found after this upload, please contact the data stewards. They can upload and link a new version on DataverseNL and add change details to a changelog.

Temporary link

It is also possible to have a publication package uploaded, but not published yet. Sometimes funders want proof of open data sharing. Inform the data stewards on time if you want to make use of this option for your publication package. Know that this link expires when the publication package is published.

Maximum Size

DataverseNL has a maximum size of 9.3 GB per file, and up to 2000 files per dataset. If any of your files exceeds these sizes, please inform the data stewards and they can help you select a different repository. This is often the case for studies with MRI data. Audio or digital video material will not be uploaded in DataverseNL. For more information on sharing MRI data please consult the [Brain MRI data sharing guide | Zenodo](#).

Contact

Willemijn Plomp & Jaap-Willem Mink, Data stewards

Mail: datastewards_psy_ped@fsw.leidenuniv.nl