

Archiving student thesis data

It is considered best practice in research to archive all research data for the sake of scientific integrity. As a student researcher you are also required to archive research data associated with a bachelor/master thesis at the university. What is considered research data, and what data you should archive is explained below.

Some study programmes at Radboud University use *RIS for Students* to archive thesis data, in which case you can find more information [here](#) on how to use it. Note that not all study programmes use *RIS for Students* for archiving, so check this with your supervisor. However, even if you will not use *RIS for Students* to archive your thesis data, this document can serve as a good guideline on what to archive.

What is considered research data?

Research data include all data that are generated during the scientific process and which are used to develop scientific conclusions. This includes both digital and non-digital information. Examples of research data include questionnaires, audio and video recordings, annotations, analysis files, notes you take, and graphs or figures.

What exactly should I archive?

Archive all data and documentation that are needed for someone (e.g., your supervisor, but also you in a few months' time) to be able to replicate your results.

***NB:** There is one type of data you **should not archive, but delete as soon as possible:** personal data which you do not need for your research conclusions, but for administrative purposes only. Think of your participants' names and the email addresses which you used to contact them. These should be deleted as soon as you don't need them anymore (e.g., after the data collection is complete) in order to protect your participants' privacy.*

The files that should ideally be archived in *RIS for Students* are listed below. Not all files may apply to your research domain. **If you have any doubts about what data to archive, you should discuss this with your supervisor, since your supervisor is the one who approves the dataset in *RIS for Students*.**

Research data files

This includes all the data files your results are based on - everything from the materials used to collect the data to the analysis files.

- Collection materials: This consists of the materials that were used to collect the data. This could be the files to run the experiment itself (including stimuli), questionnaires used, or prompts used in interviews. In a literature review, this would be the protocol used to select the literature.
- Raw data: This is the data in the form it was generated. A Qualtrics export (e.g., qsf), the responses to interview questions (audio files are preferred in some disciplines, transcripts are preferred in other disciplines), or the output from an experiment.
 - * If you did not generate the data yourself, but instead got the data from a public archive, please mention this in the readme file (see [below](#)). Include the name of the data archive and the dataset's persistent identifier (a permanent reference to an object. e.g., DOI). You then do not need to archive this raw data in *RIS for Students* since it is publicly accessible.
- (Pre-)processed data: This is the cleaned up version(s) of the data that is used for the analyses.
- Analysed data: Results of the analyses (such as tables, charts, statistical outputs). Also include the analysis scripts that were used to get the results, such as SPSS or R files.

Documentation files

It is best practice to include documentation files as they explain your data and make sure that they are understandable to an outsider or to yourself in a few months from now.

- A readme file: This is a file that provides some context for the dataset. It also lists the data files that you uploaded by name and briefly describes the content of each file. You can find an example of a readme file [here](#).
- A methodology file: A good dataset always includes a methodology file which describes the research method, but for the sake of simplicity you can upload the methodology section of your thesis.
- A codebook: If the variables or codes in your data files are not explained in the files themselves, a codebook provides an explanation of the levels of a variable (e.g., political party: republican, democrat, independent) or units for numeric values (e.g., height in cm). Sometimes this information can be embedded in the analysis tool, as is the case with SPSS files or Atlas.Ti projects.

Supplementary files

You can additionally include your data management plan if you wrote one for your research project. Further, you can include documents that are associated with conducting research with human participants.

- Ethical approval
- Blank informed consent forms
- Information sheets
- Debriefing forms