FORS
explore.understand.share.

1.0 Guide to depositing qualitative data at the FORS Data Service (DARIS)

To be realistic, making one's data ready for secondary analysis by those who did not take part in the original research project requires some time and effort. Analysts of secondary data should have at their disposal all the information needed to make effective and appropriate use of the data that you supply. This means that your data files must be clean, clear, and fully interpretable, with adequate supplementary documentation that explains your data, methods, assumptions, and also any weaknesses of your research. However, the level of effort involved in the process of documentation and preparation of data can be significantly reduced with sufficient planning and foresight. Here are a few specific ways to make the process less cumbersome:

- First, preparation should begin at project conception -- if possible, consider at the proposal stage what will be the costs of documentation in terms of time and effort. Funding agencies, including the Swiss National Science Foundation, generally allow and expect such costs to be included in grant proposals.
- Assign responsibility of the documentation process early on to a particular person who will be involved in the entire research life cycle.
- Most importantly, start the process *at the beginning* of the study, including documentation of the study purpose and design, to be integrated later into a methodology report. Don't risk waiting until the end of the research project to do documentation – chances are it will not get done!
- Consult with DARIS staff. We are available to help you in the process.

2.0 Preparing your qualititative data

The depositing of data at FORS takes place entirely online on our platform FORSbase. Before getting started with the deposit, your should prepare your data files, making sure that they are anonymised, cleaned, and documented.

Your submission of qualitative data to DARIS in FORSbase should include:

1. the data,
2. all instruments and materials used for data collection, and
3. all reports, notes, and materials on aspects relevant to data use and the research methodology.

2.1 The data

The data should be sent to us in a readable digital format. We accept audio, image, and video files in the recommended and acceptable formats shown in Exhibit 1.

Transcriptions in text format of original audio or video files should be anonymised to an adequate level, which means that it should not be easy for users to identify individuals.

Exhibit 1: Preferred and accepted formats

Category	Type of data	Preferred formats	Accepted formats
	Text	Adobe Portable Document Format PDF/A, PDF (*.pdf)MS Word (*.doc, *.docx)	OpenDocument Text (*.odt)Rich Text Format (RTF) (*.rtf)Hypertext Mark-up Language (HTML) (*.html)eXtensible Mark-up Language (XML) according to an appropriate schema (*.xml)
Qualitative or Documentation	Plain text	Unicode, Non-Unicode and ASCII (*.txt)	
	Audio*	Waveform Audio Format (*.wav)	MPEG-1 Audio Layer 3 (*.mp3)Free Lossless Audio Codec (FLAC)Audio Interchange File Format (AIFF) (*.aif, *.aiff)
Qualitative	Pictures: Raster (bitmap) images	TIFF (*.tif, *.tiff)	Adobe Portable Document Format (PDF/A, PDF) (*.pdf)JPEG (*jpeg, *.jpg)JPEG 2000 (*.jp2, *.jpx)PNG (*.png)GIF (*.gif)BMP (*.bmp)
	Pictures: Vector images	Scalable Vector Graphics SVG (*.svg)Adobe Illustrator (*.ai)	Drawing Interchange File Format DXF (*.dxf)PostScript (*.eps)
	Video	MPEG-4 (*.mpg4)	Moving Picture Experts Group MPEG-2 (*.mpg2)Motion JPEG 2000 (*.mj2)QuickTime (*.mov)Lossless AVI (*.avi)

Transcripts should have the following features:

- a unique identifier
- a document header giving brief details of the data collection event, including date, place, interviewer name and interviewee details
- a uniform layout throughout the research project
- make use of speaker tags indicating the question/answer sequence
- pseudonyms to anonymise personal identifying information
- line breaks
- page numbers

2.2 Instruments and materials

You should the following in your deposit, as applicable:

- the instrument(s) used for the collection of data (protocols, questionaires, forms), in the different languages concerned,
- any materials sent in advance to respondents (e.g., advance letters or postcards),
- any materials presented to respondents during the interviews (e.g., flashcards),

and
- any instructions or materials for use by interviewers (e.g., explanations of protocol items, frequently asked questions).

These instruments and materials should preferably be sent in standard electronic readable formats (e.g., pdf, rich text format, Word), but we may be able to accept them in paper form by mail.

2.3 Documentation, including reports, notes, and publications

Adequate documentation is especially important for reuse of qualitative data. Your submission should include any reports, papers, notes, and/or publications that provide supplementary and contextual information that can help guide users in the analysis of your data, including on:

- the context of the data: project history, objectives, and hypotheses
- the population
- sampling/selection procedures
- the interview setting
- the temporal and geographic coverage
- the structure of data files, cases, relationships between files (if applicable)
- anonymisation and data confidentiality

You should also send all signed participant consent forms. In the ideal case, you will have asked participants for their permission to use their data for research purposes. In such cases, DARIS will give access to these data to researchers, including audio, image, and video files, under strict end-user contract conditions.

3.0 Submitting your data and documentation

Once you have prepared your data, instruments, materials, and documentation, you can submit them through FORSbase. If you do not already have a FORSbase account you will need to register.

In many cases, data and documentation are delivered to DARIS at the end of a research project, and the data are generally accessible to users soon after. In some cases, however, producers of data elect to place their data under embargo, which means that while their data are preserved at DARIS, they are not accessible to users for a pre-defined period, which can be negociated in advance.