FORS
explore.understand.share.

1.0 Guide to depositing quantitative 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 Documenting and preparing your quantitative data

At a minimum, your submission to DARIS should include:

1. the data,
2. all instruments used for data collection, and
3. documents on aspects relevant to data use and the research project and its methodology (e.g., reports, papers, user guides, codebooks).

2.1 The data

The data that you send to DARIS should ideally be in SPSS format, but we can accept other formats too, provided that we can convert them. They should have at minimum the following characteristics:

- All direct identifiers should be removed, that is, variables that would allow easy identification of individuals (e.g., names, addresses, telephone numbers, etc.).
- Variables should have labels, as well as descriptions, within the data file. The descriptions can consist of shortened forms of the survey questions. Thus,

there should be clear linkages between variables in the data file and corresponding questions in the questionnaire.
- Missing values should receive an explicit code (e.g., 88 for "don't know"), and should not appear as an empty case or as a missing value attributed by default by the program (e.g., a period in SPSS).
- The data should be fully edited and cleaned. This means that before delivery, frequencies should be checked, and inconsistencies or abnormalities should be detected, repaired, or deleted (e.g., highly unlikely values).

2.2 Instruments and materials

You should transmit, as applicable:

- the instrument(s) used for the collection of data (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., question by question explanations, frequently asked questions).

These instruments and materials should be transmitted in standard electronic readable formats (e.g., pdf, rich text format, Word).

2.3 Relevant documents

Your submission should include any supplementary information that you feel can be used to help guide users in the analysis of your data, including methodology reports, users guides, codebooks, etc. While the forms of such documents may vary (e.g., sometimes as users guides, sometimes as methodology reports), whatever you provide should include the following types of information:

- the context of the data: project history, objectives, research design, and hypotheses
- population, sampling design, and sample size
- unit of analysis
- data collection mode (CATI, CAPI, mail, web, etc.)
- response rate
- temporal and geographic coverage
- structure of data files, cases, relationships between files (if applicable)
- data validation, checking, proofing, cleaning and other quality assurance procedures carried out
- information on data confidentiality, access and use conditions
- weighting
- recoded and derived variables created after collection, with code, algorithm or command file used to create them

3.0 Depositing your data and documentation

Once you have prepared your data, instruments, materials, and documentation, you can deposit them online with the help of our platform FORSbase.

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.