Archiving guidelines social and behavioral sciences

ESSB Data Coffee Breaks

October 13 & 24, 2022

Eduard Klapwijk, PhD
Research Data Steward, ESSB

Erasmus University Rotterdam

The ESSB Data Coffee Breaks

- Monthly meetings, see <u>schedule</u>

- Short, focused, and informal

- 15 minute presentations

- Twice a month

- Suggestions for topics welcome!

Archiving data for research integrity

- _Netherlands Code of Conduct for Research Integrity_ (2018)

  - Transparency as an important principle of integrity

    - Share and archive data, software codes and research materials
    - _data is open and accessible to the extent possible and remains confidential to the extent necessary_

Specific embodiment in archiving guidelines

Guideline for the archiving of academic research for faculties of behavioural and social sciences in the Netherlands (2022)

- Guidelines how to archive data, materials and information that form the basis for publications in social sciences
  - Apply to all faculty staff members, PhD candidates, research master's students
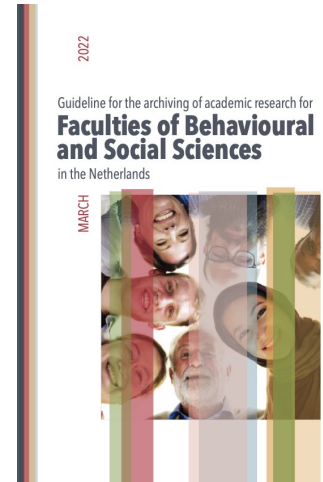
Content of the 2022 archiving guidelines

- First archiving guideline for social sciences dates from 2018
- March 2022 updated mainly with more info on qualitative data

- Main content of the guidelines:
  - Description of content of 'publication packages' (for quantitative and qualitative research)
  - Who is responsible?
  - Who has access?
  - Where to store
  - Storage period

What to include in a publication package I

1 | **A published (or accepted) manuscript or publication**

Must include a brief description of the problem definition, research design, data collection (sampling, selection and representativeness of informants) and methods used

2 | **Materials used**

Include instructions, procedures, the design of the experiment and stimulus materials (interview guide, question-naires, surveys, tests) necessary to replicate the research.

3 | **Raw Data Files**

Provide the most direct registration of behaviour or reactions of participants. Think of unfiltered export files of surveys, EEG measurements, recordings or transcripts. If needed, include all de-identification steps taken.

Pictures based on prototype of infographic currently in the making

What to include in a publication package II

Preprocessing computer code

Include code (such as Atlas.Ti/SPS-S/JASP syntax files, R-scripts, etc.) describing the steps taken to process raw data into analysis data, including brief explanations of the steps in English

4

Processed data files

Provide data (either raw or processed) that were eventually analysed when preparing the article (e.g. a data file after transforming variables, after applying selection, etc.). If the raw data was analysed directly, step 3 suffices.

5

Analysis Computer Code

Include code describing the steps taken to process the analysis data into results reported in the manuscript, including brief explanations of the steps in English

6

Pictures based on prototype of infographic currently in the making

What to include in a publication package III

Data management plan

7

Provide a copy of the most recent version of your data management plan

A readme file

8

Provide a clear readme describing who was involved in the project, when data was collected, which documents and files can be found where and how to interpret them

Ethics documentation

9

Documents related to the ethical approval (e.g. approval letter, blank consent form)

Pictures based on prototype of infographic currently in the making
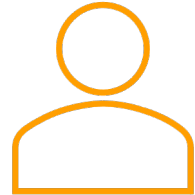
Publication packages for qualitative data

- For qualitative, interpretative methodologies:
  - Transparency possible, reproduction harder
  - Data should be archived, only documented exceptions possible

  - Additional information to record when doing fieldwork:
    - Dates of data collection
    - Type of activities (e.g., observations, conversations)
    - Interview and observation guides (if available)
    - Hard evidence of the period of time spent in the field

Who, when, and how long

- **Who?** First author is responsible for preparing and archiving the publication package
  - When collaborating with others, package at 1 university suffices
  - Supervisors of PhD + MSc (research) students bear final responsibility

- **Who has access?** First author and designated officials

- **When?** Within one month after definitive publication

- **For how long?** At least 10 years

What to share and what to archive?

Data Repository:
- Data package with curated data
- Used materials, e.g. protocols, informed consent form (NOT signed)
- Usually NO personal data
- ...
- README file

Archive:
- Raw data AND processed data
- Ethical review approval
- Informed consent forms (signed)
- Any other project documents (agreements, etc.)
- ...
- README file

Where to share?

- Discipline-specific repositories (overview at <u>re3data</u>)
  - preferably trusted repositories (e.g., CoreTrustSeal, ISO certified)
  - at least using persistent identifiers such as <u>DOI</u>

- Institutional repository
  - at EUR: the <u>EUR Data Repository</u>

- Generic data repository
  - e.g., <u>Zenodo</u>, <u>Open Science Framework</u>, <u>DANS Easy</u>

Where to archive?

- Digital archive at EUR in the making

  - Uses SURF Yoda
  - Collaboration library and research data stewards
  - Secure storage for >10 years

- Contact me for urgent cases

Yoda Portal

Some practical takeaways

- Start planning early on

- Prepare your data for archiving and re-use

- Document your data, keep it organized

- Richly describe your data

Questions

Thank you!

Research Data Steward at ESSB
datasteward@essb.eur.nl