Data Sharing and Long-Term Preservation

Astrid Recker
CESSDA Training at the Data Archive for the Social Sciences
GESIS - Leibniz Institute for the Social Sciences
@CESSDA_Data

This work is licensed under Creative Commons Namensnennung 4.0 International Lizenz.
Data Sharing
The What, Why, and How
What is data sharing (for)?

• “Data sharing is the practice of making data used for scholarly research available to other investigators” (https://en.wikipedia.org/wiki/Data_sharing)
  – creates transparency
  – enables re-use of data in new research contexts
  → thereby contributes to the efficient use of public funds (i.e. tax money)
Sharing data the end of the project

Closed
- Authorised only
- Statute
- National security data

Shared
- Company-authorised
- Corporate IP
- Specifically assigned contract
- Your personal finance records
- Combined health data anonymised

Open
- Anyone meeting the conditions
- Limited sharing license (e.g. cc-by-nc)
- A bus timetable
- Open license or waiver

Image: “The Data Spectrum”. Adapted from http://theodi.org/data-spectrum (Open Data Institute, cc-by)
Access and sharing during the research process

**Who?** Access for core research team

**What?** “Everything” is shared (working files, raw data...)

Access and sharing during the research process

Private Research Domain

**Who?** Access for core research team

**What?** “Everything” is shared (working files, raw data...)

Shared Research Domain

**Who?** Limited access for researchers outside core team (e.g. for peer review, replication)

**What?** “Stable” versions of data and documentation

# Access and sharing during the research process

<table>
<thead>
<tr>
<th>Private Research Domain</th>
<th>Shared Research Domain</th>
<th>Publication Domain</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Who?</strong> Access for core research team</td>
<td><strong>Who?</strong> Limited access for researchers outside core team (e.g. for peer review, replication)</td>
<td><strong>Who?</strong> (Long-term) access for research community or public</td>
</tr>
<tr>
<td><strong>What?</strong> “Everything” is shared (working files, raw data...)</td>
<td><strong>What?</strong> “Stable” versions of data and documentation</td>
<td><strong>What?</strong> “Final” versions of data and documentation</td>
</tr>
</tbody>
</table>

Access and sharing during the research process

Private Research Domain

Shared Research Domain

Publication Domain

Amount of shared information decreases

Amount of required metadata increases

Archives in the Research Process
We keep data safe!

www.digitalbevaring.dk
We keep data safe!

• Back-up and secure storage
• Long-term preservation
• Legal and ethical security
  – for data producers (IPR, licensing terms)
  – for participants in research (data protection)

Image: www.digitalbevaring.dk (cc-by)
We make data accessible!

www.digitalbevaring.dk

Image: www.digitalbevaring.dk (CC-by)
We make data accessible!

- Accessible formats
- Increased visibility in the world wide web (e.g. through machine-readable metadata)
- Persistent identification
- High-quality documentation

Image: www.digitalbevaring.dk (CC-by)
We offer support!
We offer support!

- We advise on
  - documentation
  - metadata
  - anonymization and data protection

Image: www.digitalbevaring.dk (CC-by)
1. Identify a suitable archive or repository

Possible selection criteria:
• data type and format
• discipline
• region or funder
• service and reputation

Search tool: re3data.org
2. Contact the archive as early as possible

Find out

• if the archive is interested in taking your data
• which criteria for submission apply
3. Discuss when, how, and in which format the data will be submitted

Clarify access conditions, legal and ethical issues.
4. Sign the Deposit Agreement

Contractual agreement clarifying rights and obligations of archive and data depositor as well as access conditions.
5. Submit the data in the agreed format at the agreed time

Possible submission channels include

• upload,
• e-mail,
• secure file sharing service (e.g. Cryptshare).
Conclusion: Sharing and Preservation

• Start thinking about this early
• Contact potential archives at the beginning of your research
• Pay specific attention to: informed consent and legal restrictions
Research Data and Self-Deposit Repositories
What is ReShare?

• Self-deposit repository for research data-open to all

• Features:
  – Part of UK Data Service, but anyone can use it!
  – Set permissions at the file-level
  – DOI minted

• UKDS staff review data collections before publishing:
  – Confidentiality and copyright
  – Documentation and file formats

• [reshare.ukdataservice.ac.uk](reshare.ukdataservice.ac.uk)
ReShare Homepage

ReShare

"Archiving and sharing research data"

ReShare is the UK Data Service's online data repository, where researchers can archive, publish and share research data, as open or safeguarded data.

Collections of data and accompanying documentation can be submitted after registering with the UK Data Service. ReShare is where ESRC grant holders submit the data from their research grants, as a contractual requirement under the ESRC research data policy. In the process related grant information can be retrieved from the RCUK Gateway to Research. We review data before their release, to ensure they conform with ethical and legal requirements.

Search and find data via the UK Data Service's Discover portal.
ReShare in Discover

Discover

Search and browse our data collections, support guides, case studies, and related publications.

Violence

Case study  Data collection  Series record  ESRC output  Support guide

ReShare (15)

Results per page: 10

Displaying 1-10 of 15 results

SN 851150 Investigating Adolescent Violence towards Parents
Rachel Condry, University of Oxford

SN 850417 Contested development?: intimate partner violence and women’s employment in urban and rural Tanzania