SUPPORTING QUALITATIVE DATA SHARING AND RE-USE: RECOMMENDATIONS FOR RESEARCHERS

BACKGROUND

Qualitative data is increasingly being used in socio-environmental systems research and related interdisciplinary efforts to address complex sustainability challenges. Qualitative data includes transcripts of recorded interviews or focus group sessions, field notes, audio and video recordings like oral history interviews, photographs, and maps, as well as policy documents, news reports, and historical archives. There are many benefits to be gained from sharing and re-using qualitative data, some of which reflect the broader push toward open science. However, although open data is increasingly becoming an expectation in many fields and methodological approaches that work on interdisciplinary topics, there remain many challenges associated the sharing and re-use of qualitative data in particular.

BENEFITS AND CHALLENGES TO QUALITATIVE DATA SHARING AND RE-USE

Broadly speaking, the benefits of qualitative data sharing and re-use fall into three categories. Scientific benefits include increasing transparency, supporting reliability and reproducibility, and providing an evidence base that can be used to scale up or down research findings. Descriptive benefits refer to the contribution that qualitative data sharing can make to characterizing and bearing witness to research contexts and subjects past and present. Finally, material benefits of qualitative data sharing include maximizing scarce research resources (both time and funding), and minimizing the burden on research subjects and communities.

Practical challenges to sharing and re-using qualitative data include the identification of appropriate infrastructure for depositing and accessing data, and the creation of standardized metadata that can provide adequate information for data re-use. In addition, epistemological approaches that guide qualitative data gathering can greatly influence the likelihood that a researcher will feel it appropriate to share that data or re-use it for synthesis purposes. The development of adequate metadata for qualitative data, including information about context of the research process, and the methodological and ethical considerations that influenced data gathering, can address many of the concerns potentially raised by the prospect of qualitative data sharing.

LEVELS OF PROCESSING AND LEVELS OF ACCESS FOR QUALITATIVE DATA

To address some of the challenges and maximize the benefits associated with qualitative data sharing and re-use, all actors across the research data life-cycle should be aware that there are many levels of access and levels of processing at which qualitative data can be shared for future re-use. Different combinations of processing and access will appropriate for different types of data and research contexts.¹

<table>
<thead>
<tr>
<th>Level of processing and definitions</th>
<th>Level of access and definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 Full text, image, etc. No redaction</td>
<td>A Data freely available for use in accordance with general use agreement of repository and standard citation practices</td>
</tr>
<tr>
<td>1 Full text, image, etc. Direct identifiers redacted</td>
<td>B Data available for use when user meets criteria set by data repository to ensure ethical data use (e.g. obtaining IRB approval)</td>
</tr>
<tr>
<td>2 Full text, image, etc. All identifiers redacted</td>
<td>C Data available for use when user is approved by the original researcher (access could depend on intended analysis)</td>
</tr>
<tr>
<td>3 Excerpt text, image, etc. All identifiers redacted</td>
<td>D Data deposit exists for archival purposes but no data are currently available (e.g. embargoed until change in sensitive situation)</td>
</tr>
<tr>
<td>4 Summary text, image, etc. All identifiers redacted</td>
<td></td>
</tr>
</tbody>
</table>

¹ Full description of these levels as well as further discussion of opportunities, challenges, and resources for qualitative data sharing and re-use can be found in a SESYNC white paper published in January 2018.
ROLE OF RESEARCHERS IN QUALITATIVE DATA SHARING AND RE-USE

Researchers from many disciplinary and institutional backgrounds are foundational to the process of generating, sharing and re-using qualitative data, and they play a key role as well in building networks and capacity for data sharing and data re-use. In the research design and data gathering process, researchers who utilize qualitative methods often spend significant time articulating the epistemological and ethical positions and safeguards that are reflected in their approach. By identifying these dimensions early in the research process, researchers create a foundation from which to make decisions about if and how to share qualitative data. Researchers must also take into account practical constraints like time and knowledge of appropriate data processing and sharing mechanisms. Finally, researchers who choose to re-use qualitative data in secondary or synthesis research projects are often looking for new sources of information to scale up or scale down their analyses, and to leverage scarce resources by drawing on existing data. Appropriate citation of others’ data products builds trust within the research community and ensures reciprocal benefits of data sharing and data re-use over time\(^1,2\).

RECOMMENDATIONS FOR RESEARCHERS

When generating new data

When qualitative and mixed methods researchers initiate new projects that will generate qualitative data, it is imperative that they plan at the outset to systematically document the data generation process and to manage the data in uniform and standardized ways. This planning and data management process should include generating metadata as quickly as possible following data gathering activities (be they in the field, in the archives or on the Internet). Researchers are encouraged to utilize data management planning and data curation tools at the beginning of a research project and to include budget lines for data management and processing in grant proposals and funding requests.\(^3\)

When depositing data

Researchers should follow the FAIR principles\(^4\) when making decisions about where, how and what data to deposit. This includes selecting an appropriate repository, making decisions about the level of access that is ethically and epistemologically appropriate, and depositing data that has been processed and includes documentation to a degree that will allow other researchers to potentially integrate the data into new analyses. Researchers are encouraged to provide data that has been processed only enough to be used appropriately by subsequent users, and to set access restrictions that are as open as is ethically possible.

When accessing and re-using data

When researchers access and re-use qualitative data for new analyses, they should provide full data citations and identify all associated content that contributes to the data re-use. Researchers are encouraged to read and engage all provided metadata to ensure appropriate and accurate interpretation of qualitative data.

---

\(^1\) For discussion of trust in open data, see Corti and Fielding (2016) and Lin and Strasser (2014).

\(^2\) For example, data sharing has been shown to be associated with article citation rates (see Piwowar et al. (2007)).

\(^3\) See the Data Management Planning Tool and Purdue University’s Data Curation Profiles as example resources.

\(^4\) For discussion of FAIR principles, see Wilkinson et al. (2016).