SUPPORTING QUALITATIVE DATA SHARING AND RE-USE: RECOMMENDATIONS FOR RESEARCH FUNDERS

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 No aggregation or analysis No metadata</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 No aggregation or analysis Idiosyncratic metadata</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 No aggregation or analysis Standardized metadata</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 Aggregation or analysis Standardized metadata</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 Analysis Summarized metadata</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 RESEARCH FUNDERS IN QUALITATIVE DATA SHARING AND RE-USE

Much like journals, publishers and repositories, public and private research funders sit at a critical nexus in the research and data ecosystem to facilitate and encourage qualitative data sharing and re-use. Public and foundation funds have supported the establishment of many common research data repositories, but the resources for long-term maintenance of these repositories vary considerably and very few focus on qualitative data\(^1\). For individual researchers, funders (especially public agencies) are increasingly setting expectations for data management and data sharing that put an emphasis on transparency, appropriate access\(^2\) and accountability of funds\(^3\). Funders are also increasingly highlighting the ability of researchers to include the costs of data curation or data access into their funding proposals\(^4\). The structure and estimation of these budget items, including allowing for costs to be charged before the end of the grant cycle but prior to all data curation tasks are complete, can greatly affect the likelihood that data are made available in a way that adheres to FAIR principles.

RECOMMENDATIONS FOR RESEARCH FUNDERS

Long-term funding for data repositories

There are many possible models for securing long-term financial stability for research data repositories, funders (both public and private) investing in research for the public good and holding their grantees accountable to open science principles have a vested interest in the creation and maintenance of repositories that reflect FAIR principles. We encourage research funders to commit to long-term funding of data repositories, especially those with infrastructure for qualitative data, and to explore investment partnerships with research institutions, journals, and publishers.

More accountability and detail in data management plans

Many funders require a data management plan as part of their grant proposal process, but ongoing assessment or accountability for these plans is highly variable, in part because of the high degree of flexibility needed across diverse funding portfolios\(^5\). We encourage funders to set clear guidelines for what constitutes data, especially qualitative data, and access within broad categories of research, and to encourage grantees to provide detail and reporting about levels of processing and levels of access as part of their data management plans and funded research.

Encourage appropriate allocation of resources for data curation and data access

Because data curation, especially for qualitative data, requires significant investments of time and possibly financial resources (as one way to maintain data repositories), it is important for researchers to include these costs in proposal budgets. Accessing secondary qualitative data also often requires significant time investment, especially when the level of access is somewhat restricted. Funders are encouraged to support budget lines for researcher time for data curation, data deposits (if relevant as a future repository funding strategy) and data access.

---

1 For a recent review of possible funding models for data repositories, see Erway and Rinehart (2016).
2 For example, the NSF Data Management Plan guidelines direct researchers to consider the “lowest level of aggregated data” that is appropriate to share for a given research community and topic.
3 A list of recent US federal agency data management policies is maintained by Northwestern University’s library.
4 For example, see the NSF’s guidelines on post-end date costs, and the NSF Sociology Program.
5 For a recent assessment of the history, application and effects of data management plans, see Metcalf (2017).

*This work was supported by the National Socio-Environmental Synthesis Center (SESYNC) under funding received from the National Science Foundation DBI-1052875.*

*The opinions expressed here are the authors’ alone and do not represent the positions of SESYNC, UMD, or the NSF.*

*Citation: SESYNC. (2018). Supporting qualitative data sharing and re-use: Recommendations for research funders. National Socio-Environmental Synthesis Center (SESYNC). Available at: [https://www.sesync.org](https://www.sesync.org).*