Five Suggestions Towards User-Centred Data Repositories in the Social Sciences

ELIAS HERMAN KRUITHOF (iD)

CHRISTOPHE VANROELEN (iD)

LAURA VAN DEN BORRE (iD)

*Author affiliations can be found in the back matter of this article

]u[ubiquity press

ABSTRACT

Data repositories for the social sciences are facing some discipline-specific challenges. This essay provides an overview of the four key challenges. In addition, five suggestions are made to strengthen domain-specific online data repositories in the social sciences, supported by good practices. Using a user-centred approach, these suggestions aim to further open up the social sciences data landscape. The essay first addresses insights regarding the sharing of quantitative and qualitative data in light of the specific needs and issues regarding these two types of data. Thereafter, we suggest (i) 'data labs' for novice data re-users, (ii) advocate for the repository as an online meeting place for researchers, (iii&iv) argue for two concrete metadata approaches for quantitative and qualitative data separately, (v) provide ideas for handling epistemological challenges of data sharing, and (vi) introduce the data repository as a possible contact medium for research participants.

CORRESPONDING AUTHOR:
Elias Kruithof

Vrije Universiteit Brussel, Departement Sociology, Brussels Institute for social and population studies, Pleinlaan 5, 1000 Brussels, Belgium

elias.herman.kruithof@vub.be

OPEN DATA IN THE SOCIAL SCIENCES

During the last decade, many social scientists have started to schedule lectures, seminars, and workshops to promote 'open science' and open data. Data stewards have been appointed to assist researchers with their data management and sharing practices, and research data repositories went online to provide a platform for research data sharing. Over recent decades, open science practices have been increasingly adopted in the social sciences. In Belgium, open data archival developments in the social sciences were largely fragmented until the launch of the Social Sciences and Digital Humanities Archive (SODHA) in October 2020. As project partners, the authors were responsible for translating the needs of social scientists to the data archival experts and software developers. In order to tailor SODHA to the needs of social scientists, we searched the relevant literature to identify common pitfalls and inspirational examples. We believe these insights may also aid others in setting up new open data initiatives and improving existing repositories. To this end, this essay provides an overview of the relevant literature regarding open data infrastructure for the social sciences as retrieved. Building on these results as well as on the feedback from our SODHA users, we address four challenges for sharing data from a social sciences perspective: (a) diverging needs of different user groups, (b) relevance-evaluation, (c) epistemological debates on the reuse of qualitative data, and (d) ethical concerns with sharing qualitative data. In the last part, we conclude with five suggestions and good practices for the design of online data repositories for the social sciences.

This dedication to data repositories has good reasons. Its scientific potential manifests itself, among other occasions, in the possibility of replicating studies, the re-purposing of old research data, and opportunities for better cooperation across academia, government, and the private sector (Engzell and Rohrer 2021; Gregory et al. 2020; Zenk-Möltgen et al. 2018). Furthermore, open social science data can be a solution for known and occurring problems, such as when only falsely significant results are published (i.e., p-hacking) or when hypotheses are made to fit certain significant patterns found in the data (known as publication bias) (Breznau 2021; Freese, Rauf and Voelkel 2022). Archived data can also enable a more profound historical perspective on social science practices (McLeod & O'Connor 2020). Data preservation prevents extra costs for data collection, which is especially relevant for hard-to-reach populations, for whom acquiring data is both difficult and costly (Corti 2007; Tarrant 2017). Finally, open data can constitute ready-to-use learning materials, offering students the experience of the complexity of data analysis in the 'real' world (Corti 2007).

We encountered the issue that many researchers do not make their data publicly available for various reasons (for an overview, see Zuiderwijk, Shinde and Jeng 2020). For instance, some worry over the potential misuse and misinterpretation of their data by unqualified users (Niu and Hedstrom 2008; Zuiderwijk et al. 2020). Others refrain from sharing because of the loss of exclusive rights to follow-up papers based on the data, or more in general because there is no good system for adequately crediting the work behind the open data (David et al., 2020; Engzell and Rohrer, 2021; Borgman 2009). Drawing from years of experience in working with linked data sources, reusing administrative data, and exploring reuse opportunities within the social sciences, we served as scholarly advisors for the Belgian repository SODHA. Data and research infrastructure is a pivotal research activity of our common research group. There exists a history spanning over 60 years in the archiving of social science data, featuring numerous well-established repositories, organizations of professionals (e.g., IASSIST and CESSDA), a periodical publication (e.g., IASSIST Quarterly), and regular annual gatherings (Downey et al. 2019). Most social scientists also have a positive attitude towards more research openness (Christensen et al. 2019; Heers et al. 2017; Jeng, He and Oh 2016; Zhu 2019). However, there is still room for the development of the open data tradition. Moreover, significant disparities exist in terms of open data practices within the social sciences. The fields of economics and psychology, in particular, have embraced a more extensive and formalized open data policy as a common practice (Freese et al. 2022). Political science has also made notable strides in this regard (Zenk-Möltgen et al. 2018). However, in sociology, the traditional perspective on open data still dominates, and well-defined open data policies remain largely absent (Christensen et al. 2019; Freese, Rauf and Voelkel 2022; Zenk-Möltgen et al. 2018). Several factors make open data less developed in the social sciences when compared to STEM disciplines (Lyon 2016; Quarati and Raffaghelli 2020; Freese et al. 2022), and social scientists are reported to be less willing to share data (Borycz et al. 2023). For example, understanding research data in the social sciences

often requires detailed contextual information, which makes the sharing process more time-consuming (Yoon and Kim 2017). Also, social scientists encounter additional difficulties due to the high ethical standards expected by the science community, the lack of funding and technical infrastructure in general, and the higher probability of the use of qualitative research data, which are often considered too complex to share and reuse (Curty 2016; Wei Jeng, He and Oh 2016; Jeng and Lyon 2016; Zhu 2019; Freese et al. 2022). Social science data repositories play a critical role in reducing researchers' perceived effort to share and reuse data, so their design can make the difference (Perrier, Blondal and MacDonald 2020; Yoon and Kim 2017).

FOUR KEY CHALLENGES FOR SHARING DATA IN THE SOCIAL SCIENCES

NEEDS OF DIFFERENT USER GROUPS

Evidently, social scientists' incentives to use secondary data first and foremost depend on how well these open data fit their information needs (Niu and Hedstrom 2008; Zuiderwijk et al. 2020). An online repository has to take all its different user groups into consideration, as they have different goals, requirements, problems, and techniques for problem solving. Experienced social scientists require well-documented data of topical fit, high quality, and comparability (Friedrich 2020). Novice researchers are heavily influenced by more experienced social science researchers when it comes to discovering, evaluating, and justifying their reuse of others' data (Faniel et al. 2012). They mainly use data that are easy to understand, and they mainly encounter challenges such as forming efficient search queries, understanding data, or opening and reading datasets. Two studies on the user profiles of social science data repositories showed a remarkably high prevalence of students: 28.6% of users of a Czech repository and 39% of users of a Finnish repository (Kudrnáčová and Trtíková 2020; Late and Kekalainen 2020).

METADATA AVAILABILITY

As mentioned above, detailed contextual information is required for understanding research data in the social sciences. Many studies on data sharing and data search in the social sciences stress the fact that colleagues, friends, professors, or supervisors are all vital sources of information about data, as briefly mentioned above (Friedrich 2020; Heers et al. 2017; Krämer et al. 2021; Yoon and Kim 2017). In our opinion, the information needed to assess relevance is too often not centralized in digital data repositories (also see Krämer et al. 2021). However, to ensuring adequate metadata for all datasets can be challenging for various reasons (see Faniel, Frank and Yakel 2019). With metadata, as with any form of communication, some kind of reductionism is inevitable. Hence, all the relevant information for secondary use is never completely foreseeable by primary researchers when they write metadata (Birnholtz and Bietz 2003). Finally, problematic data documentation is especially pronounced in the sharing of qualitative data, where the versatile and specific context of the original study is often referred to as indispensable metadata. It includes, among others, knowledge about interactions between researchers and study participants, typos in the transcripts, abbreviations, study participants' tones of voice, certain emphasis in the talk, reasons why a certain probe was made during the interview, their age, gender, social status, race, occupation, research settings, time and place, sample selection decisions, data collection methods and procedures, and complete field work protocols and transcripts (Coltart et al. 2013; Sherif 2018; Yoon 2014).

EPISTEMOLOGICAL DEBATES ON THE REUSE OF QUALITATIVE DATA

Sharing qualitative social science research data for secondary use has been widely debated (DuBois, Strait and Walsh 2018; Ruggiano and Perry 2019; Freese et al. 2022), but for different reasons than the sharing of quantitative data. The debate regarding qualitative data sharing is vastly interesting and too extensive for this essay. In short, the nature of qualitative data raises hesitations regarding data security, storage, and reuse (Alexander et al. 2020; Broom, Cheshire and Emmison 2009; DuBois et al. 2018). Epistemological debates on the reuse of qualitative data are centered around the question of whether first-hand knowledge of the original research context is vital for making sense of the nature and scope of qualitative research data or not. It is frequently heard that the knowledge resulting from processing direct experiences

is not 'transportable' either into research archives or new research contexts (Broom et al. 2009; Coltart et al. 2013). The reuse of qualitative data is still in the early stages of development, especially when compared to quantitative data (Bishop 2014; Bishop and Kuula-Lummi 2017; Coltart, Henwood and Shirani 2013; Sherif 2018). However, there are indications of an evolution towards increased open data in qualitative research (Freese et al. 2022; Mozersky et al. 2020; Mozersky et al. 2021).

ETHICAL CONCERNS WITH SHARING QUALITATIVE DATA

Besides epistemological and methodological questions, many ethical challenges are inherent to sharing qualitative data (Alexander et al. 2020). First, participant confidentiality and consent are key issues. The extent to which a participant's consent to data sharing is durable over time is an unsolved conundrum (McLeod and O'Connor 2020). Informed consent is difficult to assure when the ways in which data might be reused are unknown in advance (McLeod and O'Connor 2020). Furthermore, interview transcripts are often the intellectual property of both the interviewer and the interviewee, thus creating even more complex data sharing issues (Broom et al. 2009). Some authors advise sharing only qualitative data that can be de-identified (Mannheimer et al. 2019). However, the possibilities regarding anonymization or pseudonymization of qualitative information are often dataset-specific, as it is difficult to find a balance between losing valuable information and preserving enough information to allow reuse.

FIVE SUGGESTIONS FOR OPEN DATA IN THE SOCIAL SCIENCES

SUGGESTION 1

Based on the respective needs of novice data users and the striking prevalence of students as users of social science data repositories, pedagogical modules should be offered. We advocate more collaboration with higher education teachers who teach statistics or research methods for the development of suitable learning modules. In those modules, students could experiment in their own 'data lab' with, for example, easy-to-understand datasets and additional help in the data search process. The online 'Learning Hub' of the UK Data Service can serve as an excellent example. It encompasses dedicated pages for new users, students, and teachers; provides an introduction to widely-used data analysis software; and offers various additional resources for reusing secondary data. Additionally, there is potential for enhanced support of novice researchers by collaborating with research groups to organize workshops aimed at, for example, elucidating GDPR regulations for researchers, distinguishing between anonymization/pseudonymization techniques, conducting technical workshops (such as small-cell risk analysis and synthetic data generation), and establishing platforms to facilitate communication between data and legal experts and researchers. Online resources dedicated to the life sciences, such as the ELIXIR FAIR Cookbook[1] and RDMkit,[2] can serve as good examples.

SUGGESTION 2

Data repositories can be thoroughly improved if seen as an important meeting place for researchers, thus facilitating as much as possible the contacts between primary and secondary users, especially in the case of coping with missing metadata. Even when documented metadata are theoretically sufficient for secondary data users, some researchers might prefer actual contact with the primary researchers instead of or in addition to consulting the documentation (Niu 2009). Also, when initial trust in a dataset is lost for one reason or another, contact with data creators can help restore this trust (Yoon 2017a). The user conferences of The Austrian Social Science Data Archive can inspire other repository developers in this regard. A repository should, additionally, find a way to keep track of the experts and data creators in case they change institutions. As a good practice, refer to The Finnish Social Science Data Archive, where authors are always linked to a dataset with their ORCID.

---

1   https://faircookbook.elixir-europe.org/content/home.html.

2   https://rdmkit.elixir-europe.org/.

It also seems advisable to insert a comment section under each dataset, in which users can publicly post questions regarding the data. Such interactions could supersede or complement the common practice of 'data talk,' i.e., a casual real-life conversation among colleagues (Yoon 2017b). Contact with primary researchers is even more pertinent for the sharing of qualitative data (Broom et al. 2009). Reuse of qualitative data heavily depends on the trust reusers choose to grant to specific datasets, a process that is inherently connected with an assessment of the primary researchers' worldview, research philosophy, and experiences (Yoon 2014). Many social scientists state that prior connection to the original data and the original investigators is a condition for reuse (Yoon 2014), and the great majority of secondary studies were undertaken by researchers with at least some first-hand knowledge of the context in which the qualitative data were originally collected and analyzed (Coltart et al. 2013).

SUGGESTION 3

Repositories should aim to prevent the two main challenges with metadata: poor documentation and insufficient descriptions. Thorough documentation of quantitative data always has to be ensured on all three 'levels' of metadata (Gutmann et al. 2004; Niu and Hedstrom 2008). First, study-level metadata should outline the general description of the study as a whole, such as the purposes of the study, major conceptual categories, characteristics of the sample, researchers involved, related publications, and so on. To enhance trust in the deposited dataset, these metadata can also contain information about the publications based on this dataset and a short profile of the original researcher (e.g., expertise, home departments, main methodologies, and research interests) (Yoon 2017a). Second, file-level metadata should describe the properties of individual files such as codebooks, extensive method reports, or questionnaires. It can be helpful to include links to basic concepts explained on the Web (Koesten et al. 2021). Third, variable-level metadata should describe the measurement and coding of individual variables or groups of variables (Faniel et al. 2012). Finally, repositories should include systematic updates and news about specific datasets when they are downloaded and reused (Kern and Mathiak 2015). The dataset descriptions in the GESIS interface exemplify good practices. Per dataset, it consistently includes a concise abstract outlining the research objectives, mentions the affiliated institutions of the researchers, provides a brief overview of the dataset's variables, geographical coverage, and identifies the relevant study of data collection. Moreover, it offers a comprehensive overview of different versions and related publications per dataset.

SUGGESTION 4

Repositories should address the epistemological uncertainties of researchers when sharing their (qualitative) research data. In our opinion, repositories can provide at least two things. First, they can put forward their own clear vision that provides answers to the identified epistemological uncertainties. Then it is up to the researchers to position themselves against that vision. In addition, repositories can also provide an overview of the arguments and views in that debate, which researchers themselves can use as tools in their search for an answer to their queries. Again, the UK Data Service is an excellent illustration, offering insights into various approaches for reusing qualitative data, presenting the major debates in this field, and providing users with an extensive literature review on the reuse of such data. Secondly, it seems appropriate to provide a space in which individual researchers can express their epistemological uncertainties to the public (e.g., as part of the metadata) and where they are encouraged to justify in what manner they think these uncertainties might affect reuse.

SUGGESTION 5

Given the inherent diversity in the forms of the metadata required to describe qualitative and quantitative data, repositories should avoid requiring the use of the same metadata fields for qualitative data as for quantitative data. Metadata fields for qualitative data should provide formal flexibility along with a strong emphasis on a substantive description. For inspiration, we recommend the open-format dataset descriptions on the Qualitative Data Repository.[3] However, to provide a certain consistency in the case of in-depth interviews, a repository could ask researchers to deposit their interview guides together with indications of the dynamics of

---

3   https://qdr.syr.edu/.

the interviews (e.g., scale of openness) and a statement of the research goals (e.g., the main research questions). By solely uploading this guide, the researcher indicates which qualitative data are available without having to pass through ethics committees and without having to take the privacy of the interviewees into account.

CONCLUSION

In this essay, we identified four key challenges for data sharing and presented five recommendations for innovating domain-specific online data repositories in the social sciences to enhance data sharing. The strength of this essay lies in its usability; it offers any developer of an open data repository, at whatever scale, a thorough and contextual overview of urgent challenges for data sharing in the social sciences.

For future research, it is important to note that many fundamental aspects of data sharing in the social sciences remain understudied. Furthermore, many aspects of data sharing practices in the social sciences have not yet been sufficiently studied, e.g., which repositories do researchers choose (university repositories, public data archives, private repositories, or domain-specific repositories)? Why do researchers favor certain repositories over others? The reception of open science policies by social science researchers has also not yet been sufficiently investigated. Finally, it is crucial to understand the relative importance of informal data talks and workplace dynamics in the data sharing landscape for the social sciences.

It is time for social sciences open data repositories to fully apply a user-centered approach in the development of their online services. Such a designed online data repository can make a major difference in filling the gap between social scientists' attitudes and their actual data sharing behavior through online research data repositories. Our five suggestions can be useful on the way towards truly user-centered data repositories, the next chapter of open data.

ACKNOWLEDGEMENTS

FUNDING INFORMATION

COMPETING INTERESTS

The authors have no competing interests to declare.

AUTHOR CONTRIBUTIONS

Contributors Elias Kruithof and Laura Van den Borre designed the study. Elias Kruithof conducted the study and wrote the first draft of the manuscript. All authors edited the draft, discussed the interpretation, and approved the final version of the manuscript.

AUTHOR AFFILIATIONS

Elias Herman Kruithof orcid.org/0000-0002-8902-3246
Research Foundation Flanders, Brussels, Belgium; Vrije Universiteit Brussel, Departement Sociology, Brussels Institute for social and population studies, Brussels, Belgium

Christophe Vanroelen orcid.org/0000-0001-8619-8553
Vrije Universiteit Brussel, Departement Sociology, Brussels Institute for social and population studies, Brussels, Belgium

Laura Van den Borre orcid.org/0000-0001-5709-3533
Vrije Universiteit Brussel, Departement Sociology, Brussels Institute for social and population studies, Brussels, Belgium; Department of Epidemiology and Public Health, Sciensano, Brussels, Belgium

REFERENCES

Alexander, SM, et al. 2020. Qualitative data sharing and synthesis for sustainability science. *Nature Sustainability*, 3: 81–88. DOI: https://doi.org/10.1038/s41893-019-0434-8

Birnholtz, JP and **Bietz, MJ** 2003. Data at work: Supporting sharing in science and engineering. In: *Proceedings of the 2003 International ACM SIGGROUP Conference on Supporting Groupwork, Sanibel Island*, Florida on 9–12 November 2003, pp. 339–348. DOI: https://doi.org/10.1145/958160.958215

Bishop, L 2014. Re-using qualitative data: A little evidence, on-going issues and modest reflections. *Studia Socjologiczne*, 3(214): 167–176.

Bishop, L and **Kuula-Lummi, A** 2017. Revisiting qualitative data reuse: A decade on. *SAGE Open*, 7(1). DOI: https://doi.org/10.1177/2158244016685136

Borgman, CL 2009. Scholarship in the digital age: Information, infrastructure, and the internet. *Literary and Linguistic Computing*, 24(2): 243–245. DOI: https://doi.org/10.1093/llc/fqn041

Borycz, J, et al. 2023. Perceived benefits of open data are improving but scientists still lack resources, skills, and rewards. *Humanities and Social Sciences Communications*, 10(339). DOI: https://doi.org/10.1057/s41599-023-01831-7

Breznau, N 2021. Does sociology need open science? *Societies*, 11(1): 9. DOI: https://doi.org/10.3390/soc11010009

Broom, A, Cheshire, L, and **Emmison, M** 2009. Qualitative researchers' understandings of their practice and the implications for data archiving and sharing. *Sociology*, 43(6): 1163–1180. DOI: https://doi.org/10.1177/0038038509345704

Christensen, G, et al. 2019. *Open science practices are on the rise: The State of Social Science (3S) survey*. UC Berkeley: Center for Effective Global Action. Retrieved from https://escholarship.org/uc/item/0hx0207r. DOI: https://doi.org/10.31222/osf.io/5rksu

Coltart, C, Henwood, K, and **Shirani, F** 2013. Qualitative secondary analysis in austere times: Ethical, professional and methodological considerations. *Historical Social Research,* 38(4): 271–292. https://www.jstor.org/stable/24142699.

Corti, L 2007. Re-using archived qualitative data – where, how, why? *Archival Science*, 7: 37–54. DOI: https://doi.org/10.1007/s10502-006-9038-y

Curty, RG 2016. Factors influencing research data reuse in the social sciences: An exploratory study. *International Journal of Digital Curation*, 11(1): 96–117. DOI: https://doi.org/10.2218/ijdc.v11i1.401

David, R, et al. 2020. FAIRness literacy: the Achilles' heel of applying FAIR principles. *CODATA Data Science Journal*, 19(32): 1–11. DOI: https://doi.org/10.5334/dsj-2020-032

Downey, G, Eschenfelder, KR, and **Shankar, K** 2019. Talking about metadata labor: Social science data archives, professional data librarians, and the founding of IASSIST. In: Aspray, W (ed.), *Historical Studies in Computing, Information, and Society: Insights From the Flatiron Lectures*. Cham: Springer, pp. 83–113. DOI: https://doi.org/10.1007/978-3-030-18955-6_5

DuBois, JM, Strait, M, and **Walsh, H** 2018. Is it time to share qualitative research data? *Qualitative psychology*, 5(3): 380–393. DOI: https://doi.org/10.1037/qup0000076

Engzell, P and **Rohrer, JM** 2021. Improving social science: Lessons from the open science movement. *PS: Political Science & Politics*, 54(2): 297–300. DOI: https://doi.org/10.1017/S1049096520000967

Faniel, IM, Frank, RD and **Yakel, E** 2019. Context from the data reuser's point of view. *Journal of Documentation*, 75(6): 1274–1297. DOI: https://doi.org/10.1108/JD-08-2018-0133

Faniel, IM, Kriesberg, A and **Yakel, E** 2012. Data reuse and sensemaking among novice social scientists. *Proceedings of the ASIST Annual Meeting*, 49(1): 1–10. DOI: https://doi.org/10.1002/meet.14504901068

Freese, J, Rauf, T and **Voelkel, JG** 2022. Advances in transparency and reproducibility in the social sciences. *Social Science Research,* 107(102770). DOI: https://doi.org/10.1016/j.ssresearch.2022.102770

Friedrich, T 2020. Looking for data. Unpublished thesis (PhD), Humboldt-Universität zu Berlin, Philosophische Fakultät. DOI: http://doi.org/10.18452/22173

Gregory, KM, et al. 2020. Understanding data search as a socio-technical practice. *Journal of Information Science*, 46(4): 459–475. DOI: https://doi.org/10.1177/0165551519837182

Gutmann, M, et al. 2004. The selection, appraisal, and retention of social science data. *Data Science Journal,* 3: 209–221. DOI: https://doi.org/10.2481/dsj.3.209

Heers, M, Ferrez, E and **Morgan de Paula, E** 2017. Data sharing and re-use: Researcher practices, attitudes and needs FORS Survey of Social Science Researchers in Switzerland, FORS.

Jeng, W, He, D and **Oh, J** 2016. Toward a conceptual framework for data sharing practices in social sciences: A profile approach. *Proceedings of the Association for Information Science and Technology*, 53(1): 1–10. DOI: https://doi.org/10.1002/pra2.2016.14505301037

Jeng, W and **Lyon, L** 2016. A report of data-intensive capability, institutional support, and data management practices in social sciences. *International Journal of Digital Curation*, 11(1). DOI: https://doi.org/10.2218/ijdc.v11i1.398

Kern, D and **Mathiak, B** 2015. Are there any differences in data set retrieval compared to well-known literature retrieval? In: Kapidakis, S, Mazurek, C and Werla, M (eds.), *Research and Advanced Technology for Digital Libraries*. Cham: Springer. pp. 197–208. DOI: https://doi.org/10.1007/978-3-319-24592-8_15

Koesten, L, et al. 2021. Talking datasets – understanding data sensemaking behaviours. *International Journal of Human Computer Studies*, 146(102562). DOI: https://doi.org/10.1016/j.ijhcs.2020.102562

Krämer, T, et al. 2021. Data-seeking behaviour in the social sciences. *International Journal on Digital Libraries,* 22: 175–195. DOI: https://doi.org/10.1007/s00799-021-00303-0

Kudrnáčová, M and Trtíková, I 2020. Sustainability through the liaison with data archive users. *IASSIST Quarterly*, 44(4). DOI: https://doi.org/10.29173/iq976

Late, E and Kekäläinen, J 2020. Use and users of a social science research data archive. *PLoS ONE,* 15(8). DOI: https://doi.org/10.1371/journal.pone.0233455

Lyon, L. 2016. Transparency: The emerging third dimension of open science and open data. *LIBER Quarterly: The Journal of the Association of European Research Libraries*, 25(4): 153–171. DOI: https://doi.org/10.18352/lq.10113

Mannheimer, S, et al. 2019. Qualitative data sharing: Data repositories and academic libraries as key partners in addressing challenges. *American Behavioral Scientist,* 63(5): 643–664. DOI: https://doi.org/10.1177/0002764218784991

McLeod, J and O'Connor, K 2020. Ethics, archives and data sharing in qualitative research. *Educational Philosophy and Theory*, 53(5): 523–535. DOI: https://doi.org/10.1080/00131857.2020.1805310

Mozersky, J, et al. 2020. Research participant views regarding qualitative data sharing. *Ethics & Human Research,* 42(2): 13–17. DOI: https://doi.org/10.1002/eahr.500044

Mozersky, J, et al. 2021. Barriers and facilitators to qualitative data sharing in the United States: A survey of qualitative researchers. *PloS ONE*, 16(12). DOI: https://doi.org/10.1371/journal.pone.0261719

Niu, J 2009. Overcoming inadequate documentation. In: *Proceedings of the ASIST Annual Meeting.* Vol. 46. Silver Spring, MD: American Society for Information Science and Technology. pp. 1–11. DOI: https://doi.org/10.1002/meet.2009.145046024

Niu, J and Hedstrom, M 2008. Documentation Evaluation Model for Social Science Data. In: *Proceedings of the ASIST Annual Meeting*. Vol. 45. Silver Spring, MD: American Society for Information Science and Technology. p. 11. DOI: https://doi.org/10.1002/meet.2008.1450450223

Perrier, L, Blondal, E and MacDonald, H 2020. The views, perspectives, and experiences of academic researchers with data sharing and reuse: A meta-synthesis. In: Dorta-González, P. (ed.), *PLoS ONE,* 15(2). DOI: https://doi.org/10.1371/journal.pone.0229182

Quarati, A and Raffaghelli, JE 2020. Do researchers use open research data? Exploring the relationships between usage trends and metadata quality across scientific disciplines from the figshare case. *Journal of Information Science,* 48(4): 423–448. DOI: https://doi.org/10.1177/0165551520961048

Ruggiano, N and Perry, TE 2019. Conducting secondary analysis of qualitative data: Should we, can we, and how? *Qualitative Social Work*, 18(1): 81–97. DOI: https://doi.org/10.1177/1473325017700701

Sherif, V 2018. Evaluating preexisting qualitative research data for secondary analysis. *Forum, Qualitative Social Research/Forum, Qualitative Sozialforschung,* 19(2). DOI: https://doi.org/10.17169/fqs-19.2.2821

Tarrant, A 2017. Getting out of the swamp? Methodological reflections on using qualitative secondary analysis to develop research design. *International Journal of Social Research Methodology,* 20(6). DOI: https://doi.org/10.1080/13645579.2016.1257678

Yoon, A 2014. 'Making a square fit into a circle': Researchers' experiences reusing qualitative data. *Proceedings of the ASIST Annual Meeting*, 51(1). DOI: https://doi.org/10.1002/meet.2014.14505101140

Yoon, A 2017a. Data reusers' trust development. *Journal of the Association for Information Science and Technology*, 68(4). DOI: https://doi.org/10.1002/asi.23730

Yoon, A 2017b. Role of communication in data reuse. *Proceedings of the Association for Information Science and Technology*, 54(1). DOI: https://doi.org/10.1002/pra2.2017.14505401050

Yoon, A and Kim, Y 2017. Social scientists' data reuse behaviors: Exploring the roles of attitudinal beliefs, attitudes, norms, and data repositories. *Library & Information Science Research,* 39(3). DOI: https://doi.org/10.1016/j.lisr.2017.07.008

Zenk-Möltgen, W, et al. 2018. Factors influencing the data sharing behavior of researchers in sociology and political science. *Journal of Documentation*, 74(5): 1053–1073. DOI: https://doi.org/10.1108/JD-09-2017-0126

Zhu, Y 2019. Open-access policy and data-sharing practice in UK academia. *Journal of Information Science,* 46(1): 41–52. DOI: https://doi.org/10.1177/0165551518823174

Zuiderwijk, A, Shinde, R and Jeng, W 2020. What drives and inhibits researchers to share and use open research data? A systematic literature review to analyze factors influencing open research data adoption. *PLoS ONE,* 15(9). DOI: https://doi.org/10.1371/journal.pone.0239283