

Big Social Researchers Interview Guide

Project Overview

Project Title

Connecting communities of practice: Using strategies from qualitative data curation to support big social research

Research summary

Big social data (such as social media and blogs) and archived qualitative data (such as interview transcripts, field notebooks, and diaries) are similar, but their respective communities of practice are under-connected. Research with both types of data repurpose existing social data to advance discoveries in social science. However, despite these similarities, big social research has not yet been widely framed as a form of qualitative data reuse, and qualitative data reuse has only begun to be discussed through a big social data lens. Qualitative data reuse is a more established practice, and therefore has more developed data curation strategies to support data sharing. My research investigates how data curation practices from each of these communities can inform the other for mutual benefit. The research will use interviews of qualitative researchers, big social data researchers, and data curators to gain insights into different community approaches to research and data sharing.

Research background

This research asks: how can data curators best handle qualitative and big social data to support ethical, epistemological, and legal data sharing practices?

My review of the literature revealed that there are six key issues that pose challenges for both groups. During the interview, I will ask questions about your personal experience in each of these six topic areas, plus introductory and wrap up questions. The interview will take 60-75 minutes.

Thank you for taking the time to speak with me! Your interview will help to improve data curation practices across disciplines.

Interview questions

We'll start recording the interview now.

Introductory question

Tell me about the type of research you do and what kind of data you produce.

Identifying a specific example

Please describe a recent time when you:

- collected big social data for research; or
- reused big social data that was collected and shared by someone else; or
- prepared big social data for publication or sharing.

Please also describe your data collection method (API, scraping, shared dataset, etc.)

Was this example part of a grant-funded project that required specific treatment of the data?

Did you have a data management plan?

Did you publish any of the data from your example?

- Is the data published in a repository? Which one?
- What are your plans for storing, retaining, and deleting data in the future?
- Who has access to the data?

Context

Halavais (2015) suggests that “when we collect data from [social media] platforms (just as when we collected data in traditional spaces), context matters.” However, the context of a social media post may be absent or difficult to understand. Social media posts are by nature short pieces of text, images, videos, etc, taken from a larger context of personal and public life. This out-of-context effect is only compounded when data are amassed at a large scale.

1. Tell me about a time (if any) during the process of your example when you considered the issue of maintaining and understanding the data's context (i.e. contextual information about the community where the data was collected, contextual information about respondents)?

- What were your considerations? What concerns did you have? What factors helped you better understand/communicate context? What factors prevented you from understanding/communicating context? Was your research affected by incomplete contextual information?
- Did you consult with anyone, consider other research projects, or refer to literature, policies, or guidelines regarding this issue? Please explain.
- What strategies did you use to discern/understand context during your example?
- If the example includes publishing your own big social data, what strategies did you use to communicate context to future users?

Data quality

2. Tell me about a time (if any) during the process of your example when you considered the issue of data quality (for example, missing data, bots, bias, quality of method)?

- How did you assess quality? What data quality issues arose?
- What data quality concerns did you have?
- What factors helped you better consider data quality issues? What factors prevented you from considering data quality?
- Did you consult with anyone, consider other research projects, or refer to literature, policies, or guidelines regarding this issue? Please explain.
- What strategies did you use (or did you see used) to communicate, describe, or clarify data quality issues in your example? Please describe in detail how those strategies helped you.

Data comparability

3. During your example, did you compare and/or combine multiple big social datasets?

- or: Did you consider comparability or interoperability of your dataset?
- Did you consult with anyone, consider other research projects, or refer to literature, policies, or guidelines regarding this issue? Please explain.

If no:

- why not?

If yes:

- Why (for what purpose) did you combine the datasets? How did this advance your research?

- what dataset did you combine it with - where did that data come from? Your own or someone else's?
- What strategies did you use to combine multiple qualitative datasets?
 - what challenges did you encounter and how did you address them?

Informed consent

4. Tell me about a time (if any) during the process of your example when you considered informed consent.

- Did you consult with anyone, consider other research projects, or refer to literature, policies, or guidelines regarding this issue (including IRB)? Please explain.
- Have you used any other type of consent besides informed consent per se (e.g. broad consent, focus groups, community advisory boards)
- Did you feel that participants in your research would expect to give informed consent for the research?

Privacy and confidentiality

5. Tell me about a time (if any) during the process of your example when you considered issues of privacy (e.g. protecting data during research, considering restricted access or TweetIDs only if publishing).

- Did you consult with anyone, consider other research projects, or refer to literature, policies, or guidelines regarding this issue? Please explain.
- During your example, what do you think the participants' expectations of privacy were?
- Did you feel that you had to make any compromises about participant privacy in order to conduct your research?
- What strategies did you use to address the issue of privacy?

Intellectual property

6. Tell me about a time (if any) during the process of your example when you considered intellectual property concerns when you conducted your research and/or published your data (e.g. social media platform terms of service, participant intellectual property).

- Did you consult with anyone, consider other research projects, or refer to literature, policies, or guidelines regarding this issue? Please explain.
- Do you consider your research to fall under Fair Use?

Additional issues

7. Are there issues or challenges that arose during your example that I haven't asked you about?

8. Who else should I interview? I am trying to reach big social researchers, qualitative researchers who have published or reused data, and data curators who have worked with qualitative or big social data.