

Data Interview Guide
Angela Okune
May 2019

Thank you for agreeing to be interviewed, I really appreciate your time. I'll do my best to keep this to an hour in order to respect your time.

[Review the informed consent document together and obtain signature and any specific requests for data storage.]

I want to review an example data artifact that you kindly shared with me earlier this year. I uploaded this to a private group on a platform that I've set up called Research Data Share as part of my dissertation research. The reason I uploaded it was so that we could look at it together in this new format to think and talk about what might concern you and your organization about sharing such a piece of data. Please be assured that currently, this data is completely private and only visible to me and the server where this platform is hosted was set-up and paid for me. Again, the intent is to have a new format to view this data and to begin to imagine what worries and opportunities might be for this type of data sharing platform.

Access & Sharing

Would you feel comfortable sharing this data?

- ... no, not with anyone.
- .. yes, with immediate collaborators including clients.
- ... yes, with others at this organization.
- ... yes, with others in my field of study.
- ... yes, with others outside of my field.
- ... yes, with anyone.

Who else would need to be part of the decision making about whether or not to share this data? (management, funders, participants?)

When would you be comfortable sharing this data?

- Immediately after the results are published/project is finished.
- 6 months after the project is completed.
- 12 months after the project is completed.
- Other

Who do you think would be interested in this data? (e.g. other researchers in my field, those who were studied, policy makers, private companies, donors, etc.)

How would you imagine this data being used by such people? What value would this data have for them?

Would you place any conditions on sharing your data with the groups? (e.g. requiring some form of acknowledgement, money, etc?)

If you were on the user side, would you use this data? Why/why not? How?

What additional information would you need to have about this data in order to use it (better)?

How do you imagine people would find this data? (e.g. google search; referred to it by someone; etc.)

On Ontology

How would you amend the title of this dataset?

How would you amend the description (e.g. detailed annotations, a codebook, a data dictionary, etc.)?

What categories or keywords would you tag the data with to make it easier for other researchers to find?

Do you think the organization and description currently here is sufficient for another person with similar expertise to be able to understand and properly use the data?

On RDS

What do you think have to be important components for this data sharing platform?

- Being able to cite the dataset (proper attribution with a DOI)
- Restricting access to the dataset to authorized individuals only
- Ensuring that someone accesses the entire dataset and all corresponding materials together so that they don't take one data point out of context
- Being able to see usage statistics on how many people have accessed this data
- Being able to gather information about the people who have accessed or made use of the data

Beyond usage statistics are there any other measurements or analytics that you would like to apply to your data?

Would you want to submit this data into the repository yourself as part of a manual process? Or would you want the process to be automated through a trigger (e.g. a date, an event, etc.) without your intervention?

On Assessing Quality of Dataset

- How do you evaluate how good or clean a particular dataset is? What makes you flag a data-set that seems suspicious or like "bad" data?
- What do you believe a researcher's responsibilities should be when s/he uploads data?
- What do you think should be the responsibilities for anyone who accesses this data?

IP & Ownership

- What was the funding source that enabled you to do the research that produced this data?
- Do any of your funding sources require you ...
 - o to draft a data management plan beforehand?
 - o To share your data with others, publish your data, deposit your data into a data repository?
 - o Preserve your data beyond the life of the funding?
- Is the dataset bound by any privacy or confidentiality concerns?

Data Discovery and Use

Have you ever looked for datasets created by others for your research?

What kinds of data sets have you found? Where? Did you use it for your research? How?

What datasets have been particularly hard to find?

[Why do you think that is?]

Conclusion

Now that you've seen this, what do you think could be a benefit of digitally archiving such data?

...a risk to digitally archiving data?

- What do you think others in your organization would say about sharing data?

Thank you so much for your time!

[If they approved for the data to be shared in PECE:] I will share the resulting audio and transcript with you by email for your approval before posting.

Guide for myself for post interview reflections and fieldnotes:

- Did the respondents use of the term “data” change throughout the duration of the interview? Did their use of “open” change over time?
- What other concepts did they rely on heavily throughout the interview? What terms did they use the most?
- Are there people that were mentioned that I should follow-up with? Organizations? Events (future or past)? Places? Laws? Funders? Technologies?
- How did this person characterize the data ecosystem/infrastructure/communities in Nairobi? What adjectives were used? Were they optimistic/angry/what kind of emotional response to the discussion?
- Anything unusual from the other interviews / noteworthy? Any particular triggering questions? Any additional questions I should add or remove?
- What kind of training did this person have?
- How does the physical geography and built environment of Nairobi influence the movement of data?