

Indexing Data Types

Method activities	Data type (pdfs? mp4s? jpegs? tabular?)	Preservation (where will this data live?)	Access & security concerns (Who should and shouldn't see?)	Discoverability (How should and could this data be made discoverable?)
Interviews	m4a audio files; Word docs of transcriptions; pieces of paper with sketches that interviewees did while talking (and JPEG photos of these)	My hard drive, per my IRB	Much of it is not anonymized so no one else should see it at this point	Anonymization would be difficult, because many of these interviews are scientists talking about their research in detail to an extent that they are identifiable. But maybe smaller excerpts could be compiled together along certain themes, and these themes could serve as vehicles for discovery.
Participant observation	Word docs with fieldnotes; physical ephemera and JPEG photos of some of these; spreadsheets; JPEGs and mp4s; physical books	My hard drive, per my IRB	Much of it is not anonymized so no one else should see it at this point	A more processed version of fieldnotes could be made public and searchable within some kind of online system—but this would take a lot of time to anonymize adequately, remove more personal information of my own, etc. A better way to do this might be sharing the spreadsheet I maintained that has various encounters/artifacts/narratives around sugarcane. It's a lot more processed than my fieldnotes but still "rawer" than any written work. The question of what degree of "rawness" makes sense for sharing data is really interesting to me, because of course none of our data is raw to begin with.
Online archival research	PDFs, JPEGs, screenshots	My hard drive	Anyone could see this; it was mostly publicly available to begin with	This could be hosted on a website somewhere, but it would be a ton of work to upload and catalog all of these artifacts. Right now they exist in organized folders on my computer and I took notes on them as I found them. I'm not sure I have the time as a grad student to go back now and prepare them for online sharing, but if I had done this from the beginning I might have been able to. A lot of the interesting info I pulled from these artifacts are now in my spreadsheet mentioned above, so again that spreadsheet might be a more feasible/practical route for sharing some of this information in a more processed way.