QR04\_transcript\_deidentified

**SUMMARY KEYWORDS**

data, people, ethnographic, context, archives, community, consent, question, qualitative data, researcher, anthropologists, research, studies, report, read, record, reuse, practices, problems

**SPEAKERS**

Sara Mannheimer, QR04

**Sara Mannheimer** 00:39

Okay, so basically, the research that I'm doing to give you a little more information. My idea is that archived qualitative data, and big social data are similar. It's like people, they're narrative pieces of people's thoughts that are available for use for other purposes. So I'm thinking that reusing qualitative data is similar to just using social media data, you're looking at data that sort of like in a bigger context, and you're reusing it for new purposes. And so this project is about learning more about qualitative data reuse, and learning more about big social data research. In order to try to... for instance, qualitative data curation and reuse is more established, and trying to learn from your community and sort of draw connections with the social media research community to build data curation practices that will lead to more responsible research. So that's the plan. So I reviewed the literature and I identified six key issues that pose challenges for both qualitative data and big social data. So the interview is structured around those six topic areas. We have context, data quality, data comparability, informed consent, privacy, and intellectual property. So I'll ask you one question at the beginning, and then just a wrap up question at the end. So it should take about 75 minutes, max, I think. Okay. Any questions before we start? Okay, great. So can you just first tell me about the type of research that you do and the type of data that you create or use?

**QR04** 02:42

Yeah, so I'm an anthropologist, I'm a mixed methods researcher. Most of the research that I do has a qualitative—mainly qualitative data collection component, and it becomes mixed methods, because some of that is analyzed qualitatively using thematic analysis or coding, and some of the coding can be converted into quantitative data. So most of my data sets begin as qualitative data. Um, yeah, I do ethnographic fieldwork, I do a lot of cross cultural fieldwork. So implementing the same or similar data collection protocols across a number of sites simultaneously. I do some larger end research sometimes. And most of my experience with secondary data analysis is working in [an ethnographic data archive]. And the purpose of those analyses is basically just to figure out how a cultural practice manifests in the ethnographic record, which is [many decades] deep.

**Sara Mannheimer** 03:52

Great. So for this interview, I was thinking we could identify a specific time when you either prepared your own data for publication or sharing or reused existing data yourself. And then we can sort of like go between different examples if we want, but it's nice to have least have one so we have something specific to draw from. So can you think of in a recent example, a somewhat recent time when you either published data or reused qualitative data?

**QR04** 04:27

Yes, I can think of a paper in particular in which I did quite an extensive analysis of [data from the ethnographic data archive].

**Sara Mannheimer** 04:35

Great. Tell me a little more about it.

**QR04** 04:38

Sort of the theoretical question?

**Sara Mannheimer** 04:40

Um, yeah, yeah, I guess just to give us like a foundation so I can understand as we go through the rest of the question.

**QR04** 04:48

Yeah. So I work on [a specific resource scarcity]. And this is a phenomenon that had not been studied very much in anthropology at the time when I started working on it about 20 years ago. And so it was, unfortunately necessary to read the entire ethnographic record on this. And so the good thing about [the ethnographic data archive] is that the data are already coded. So you can run searches, and it pulls up pre coded data. So you're not you're not working, you know, with [many decades] of full-on analyzed texts, it's coded at about the paragraph level. So, you know, you might run a query on a culture group. And, you know, you get 60 paragraphs, which is a lot more manageable. Yeah, considering how much data there is.

**Sara Mannheimer** 05:48

Yes, yeah. Okay. And so was this paper that you wrote part of a grant funded project where you had had specific requirements for managing the data?

**QR04** 06:02

Ma'am, no! I did this [grant proposal to a national funding agency] that I felt was this amazing idea. It was a rejected proposal. And so I was like, that's fine. This is now a lit review. And it's going to be a major theoretical piece, which it did turn out to be basically, you know, after my sweet baby proposal, which was my first big proposal that I wrote, on the tenure track, was bitterly, repeatedly rejected, I decided to do this secondary analysis.

**Sara Mannheimer** 06:32

Nice. Okay. So you didn't have like a data management plan or anything like that as you went through? And I guess with secondary analysis, maybe you weren't downloading anything at all right? You were just querying what was in the database?

**QR04** 06:45

That's right. So everything there is already posted. Okay. I can speak a little bit to my own experiences with my data, if that's useful.

**Sara Mannheimer** 06:58

No, actually, that's okay. I have talked to other people who have that. So since your example is data reuse, that's unique, so I want to stick to that. Cool. So let's start with context. So I have a quote, I'm gonna read you to help. I mean, you understand what it is, but sort of to help you come to let you know where I'm coming from. Qualitative research is a process that may include deep and prolonged contact and connection with research subjects, attempting to understand the subjects within their own context. So qualitative data are highly context dependent. And context is a source of data, meaning, and understanding in itself, ignoring context, underusing it, or not recognizing one's own context driven perspective, could result in incomplete or missed meaning and a misunderstanding of human phenomena. So that's kind of what I mean when I'm talking about context. And so can you... With that in mind, can you tell me a time if any, during this project, when you considered the issue of understanding, maintaining or communicating the data as context? And then how, what your considerations were and how you dealt with them?

**QR04** 08:05

I have many thoughts. Let me begin from the perspective of somebody who generates a lot of data, right? I think that the problem of having a decontextualized interview to me is, is a bit worse than having decontextualized ethnographic data, because the the purpose of ethnographic data is to document the context, or to explore a question in context. So you can access more contextual data in secondary ethnographic data than you would be able to an interview data. Right. So that's helpful. And the second answer I have is that I think, you know, when you're querying data in this way, you need to have a research question that's appropriate to what the data can tell you. And you need to confine the conclusions. You draw very, very strictly and carefully to what the data can and can't tell you. So one of the things that I was looking for in these data was historic coping strategies, to deal with [the resource scarcity that I was studying]. And using previous findings from similar resource scarcities to see if there was any ethnographic evidence of people doing these things. So let's say I had, you know, five domains and five coping strategies in each of the domains and I'm looking you know... So here's an example. Um, you know, there's evidence that in famines, people will migrate to places that have better food access, or they might send one family member to go engage in wage labor and send back money to increase food buying power, or they can foster children so they, they can send the whole family... they can send to labor, or they can send the most vulnerable members, the children, to be fed elsewhere. But the more vulnerable adults stay right. So There is no documentation of child fostering around [the resource scarcity] whatsoever in the theoretical literature, right? So I'm diving very deep into these ethnographies to see you know, does anybody mention it offhand? This is where context is important. That was not the topic of any research study in 150 years in ethnography, but sometimes people mention it in an offhand way, right? So if it was never mentioned, I can't say it didn't happen, I can only say there was no evidence present in the ethnographic record. Right. An example of, you know, asking the question very carefully, and drawing very careful conclusions. And then, another thing about context is that, um, so there's a very well documented history of racism in ethnography, and colonial foundations of ethnography. And there's a well documented understanding that the sort of colonial apparatus used [redacted] policy as a way to bring people under colonial control. [Details redacted]. However, the depictions of [some practices that are adjacent to this resource scarcity] in the ethnographic record are very racist.

**Sara Mannheimer** 11:55

Yeah.

**QR04** 11:56

So it'd be an example of where that context is present. Yeah, it creates some challenges in interpreting the secondary data for the analysis. And even, you know, not even just interpretive challenges in terms of data quality, but also ethical challenges, you know. What does it mean to use these data? You know, how trustworthy is the observer?

**Sara Mannheimer** 12:24

Yeah, how did you, like, deal with those? How did you address them? And did you use that data? Did you document that in your paper?

**QR04** 13:35

Yeah, it was not documented in the paper. But these are, this is a well documented problem, both around historical ethnography and around the [ethnographic data archive] in particular. So it's a known issue. Um, I would say that most of that stuff was not... So as I mentioned, there were certain things I was looking for. And they were in the context around what I was looking for. And were not necessarily around the specific phenomenon I was studying, right. So they weren't necessarily reportable. They were contextual. Um, I would say I'm still grappling with how to bring that into my teaching. And it certainly informs how I teach. It informs how I teach these methods and how to work with secondary data. And it's still informed, right, I'm still reflecting on, you know, what is the most ethical way to engage with these data? And there are also consenting issues in the historical record. So for example, Indigenous societies for whom sacred or secret data are reported, in which there were not what we would consider anything remotely like appropriate consenting procedures today. And so there are real ethical quandary is around the reporting of those data.

**Sara Mannheimer** 13:56

Like, should you use it at all again?

**QR04** 13:59

That's right.

**Sara Mannheimer** 14:01

So from like a data curation perspective, I'm curious about whether there's like, whether there's documentation in the [ethnographic data archive] about these things. So like, are certain records flagged? Or is there like, additional documentation of the context? You know, metadata tags, or, you know, studies that are drawn from the data that you can link to that have more information about the methods? How, let's start with context, I guess, but like, how did how was that communicated to you when you were trying to use the data?

**QR04** 14:40

So I'm, I'm unaware of those flags being in there. But it's also true that I did not look for them. So it's possible that they're there and I wasn't looking for them.

**Sara Mannheimer** 14:52

I feel like if they were there, where you would have seen them, you know,

**QR04** 14:56

They were not highly present such that I did use the data without running into them, I have had some conversations with [people involved in managing the ethnographic archive], you know, informally advocating for engaging with some of these issues more aggressively so that when we bring early career and junior scholars into interaction with these data, they are having more of the experience that you've just asked about, right? They were aware of these things are in that, you know, there were certainly times when I was reading the data, and I would have to, you know, get up and walk around and kind of take a breather, because it was disturbing.

**Sara Mannheimer** 15:41

Does it tell you who the researchers are who gathered the data?

**QR04** 15:45

Yeah.

**Sara Mannheimer** 15:45

Okay. And where they gathered the data, like some basic information about each of the studies?

**QR04** 15:53

Yes.

**Sara Mannheimer** 15:54

Okay. And did you feel that was enough? Like, did you What more would you have wanted context wise?

**QR04** 16:03

For me, you know, I've been doing this for decades, I understand what the basic terrain is, I, you know, it... reading it can be shocking, but I'm aware that it's there. Right. Yeah. And so for me, it, it was what I expected on some level. Um, the concern that I have is for junior scholars who may be, you know, first of all, might not have that foundation to expect it, you know, the, the archives are used in teaching and undergraduate classes, and an undergraduate would not have the grounding, um, to go into the data expecting to see that. And I also think, you know, it's quite likely that people are from the societies that are depicted in the data. And I think it could be particularly traumatic for students and early career scholars from the societies to see their societies depicted in that way.

**Sara Mannheimer** 16:55

Yeah.

**QR04** 16:55

Those are concerns that I have.

**Sara Mannheimer** 16:58

Oh, wow, complicated. And challenging. Okay.

**QR04** 17:03

I mean, I know anthropology very well, and we have these issues, but I assume that, you know, all social sciences have some of these, and it's where, you know, people who had access to generating scholarship 100 or 150 years ago, do not share the ethical commitments that we have today.

**Sara Mannheimer** 17:20

Yeah, I mean, moving away from ethical questions, even just to like, where those researchers were coming from? Like, do you think that this data, I mean, is like trustworthy? You know, like, if those researchers were coming into communities from completely other contexts, with different ideas of what that meant, you know, like, how trustworthy is the data they collect?

**QR04** 17:51

This is a very widely discussed issue in anthropology and start, I mean, it's, there's no argument that some of the data is inaccurate, it's well established that, you know, there's a great deal of inaccuracy in the data, right. But it's also true, that the data has some value, right? So some, the data is very, you know, a lot of the early researchers were just writing down everything people said, right, and so, you know, they're just trying to have songs dictated to them. And so there are things that you know, so one of the things that I've done in my research is to look at those data and see, how do people use ritual to address [resource scarcity] related anxieties, right? So anxieties around [related challenges], we're going in and looking at, you know, a directly transcribed song, or a prayer has some value, right and inaccuracies, but there's a value to it.

**Sara Mannheimer** 18:55

How else? I'm especially interested in like strategies that you use as a researcher to sort of solve these problems. And then potentially that data curators could, like learn from you know, so when you're thinking about those inaccuracies? Are these problems with like, data quality or like trustworthiness of the data? What, what better informed you to understand like the level of quality? And like, what did you use to sort of analyze or assess the quality? You know what I mean?

**QR04** 19:31

I do, yeah. So, um, okay, let's start with what I'm thinking when I'm looking at these data and then pull back from there. Um, when I'm looking at the data, I'm looking for that ethnographers report of what they directly observed, or like, walking next to someone and see a pit and there's water in the pit and this is what we do, right. So it's a detailed first person, step by step. observational, interactive. Report of what occurred, right. So that kind of situation. Yeah, there are still perceptual issues. But it's much less so than a sort of like, a third hand gossip reports, like someone that this might have happened, right. And so that, to me is an important element of data quality. This is where having the context versus just a very thin report can be quite meaningful, because if you're looking at how they're depicting a range of issues, particularly sensitive issues, you kind of get a sense for the positionality and the ethical commitments of the researcher. Right. Oh, how hard are they trying to report what they've seen outside of their own meaning making systems? Right? Yeah. So adopting labels, rather than thick description is sort of a flag for that, right?

**Sara Mannheimer** 21:00

A flag to indicate lower quality?

**QR04** 21:04

yeah, so I'm, right. Example: in [a country], there's a—just because we talked about child fostering—there's a child fostering system, right. And it can be variously described as a kinship system, anything ranging from a kinship system to a child slavery system, right? And it depends on on what pieces you're encountering. And it depends on your positionality. Right? It depends on your commitment to contemporary child labor laws, right? So you can interpret what you see in the context of, you know, contemporary child labor laws, and it looks a certain way. Or you can dive down into thick description and say, you know, what labor was the child doing? How did the child come to be here? Were there any payments? You know, was there any exchange and then you can allow the reader to draw conclusions about, you know, is it a kinship system? Or is a child slavery system?

**Sara Mannheimer** 21:58

Yeah. Okay, that makes a lot of sense.

**QR04** 22:00

Can you zoom all the way in and give a detailed description and allow the person to to interpret what they've seen in the framework that they're analyzing it in, or are you just trying to slap a name on it?

**Sara Mannheimer** 22:11

Yep. Okay. Nice.

**QR04** 22:14

All right, names get slapped on. It's kind of like, Oh, I don't know about that.

**Sara Mannheimer** 22:17

Yeah. Right. Okay. Like, just tell me what you saw. Yeah. Okay. Right.

**QR04**

So to me, so again, like with the [redacted] example, like, if you see the term [a term that is currently understood to be pejorative], it's like, I don't I don't care what you think about the [meaning of it]. I just want you to tell me, what was the... How was [an activity conducted]? Was it [details redacted]?

**Sara Mannheimer**

Yeah. Yeah. Um, have you encountered any other data quality issues? Like, problems with missing data? or?

**QR04** 22:53

Yeah, nobody wants to talk about [the resource I study]. I mean, you know, it's like the most uninteresting thing. And so the data is almost always missing, right. And when we get reports, when it's scarcity is noteworthy. So day to day practices, you know, you will only see the reports largely when when it deviates from expectations. So that means that the vast majority of reports [about this resource] in ethnographic record are missing. And so I have to be very careful about the conclusions that I'm drawing. Right.

**Sara Mannheimer** 23:22

Yeah. Okay.

**QR04** 23:26

So there was a part of that, because I started with the how I'm looking at the data, but I wanted to go back and say, Oh, yeah, um, because of the way that anthropologists make contributions in [subdiscipline] in particular, there is often you know, especially with more known ethnographers, if there are problems with their data or interpretation, those are usually documented in the scholarly record. So those flags could be put on, right.

**Sara Mannheimer** 23:56

Mm hmm. Yeah. And then yeah, I mean, that's part of what I'm trying to see is like, how much are those flags put on now? And how much do they help you as a researcher? If they are there, you know? Yeah. Okay. Great. Well, let's move to data comparability. Did you compare it sounds like you did compare and combine multiple data sets? And why did you do it? How did it advance your research? And then what challenges did you encounter?

**QR04** 24:31

Yeah, so um, you know, all of my work is fundamentally cross cultural. Anthropologists are trying to understand, you know, how do humans you know, broad human patterning, generally speaking. So when I collect my own data, it's usually cross cultural in nature. Um, and I have kind of strong feelings about that because of what we've already discussed with the gaps in the observation. And interview data, you can't really compare answers unless you've asked people the same thing. And you've made some kind of effort to elicit comparable data. Yeah. And so this is where when you're working with secondary data and the ethnographic record, you're never getting any sort of, you know, comparable elicitation. And so, you know, this means that the kinds of questions you would ask of these data really have to be profoundly exploratory. And they need to be fairly novel in the sense that, you know, if it's not a new question, probably there's better documentation in the scholarly literature.

**Sara Mannheimer** 25:41

Yeah. Um, so when you're going, sort of just like, practically as you went through it, did you use certain search terms, and then just read carefully into each ethnographical report? I'm trying to think like, what other challenges were there? You know, you created the right research question to help with this comparability issue. But what other challenges did you encounter when you were trying to like, compare a bunch of different ethnographers from decades to one another?

**QR04** 26:24

It's just these gaps, you know? Huge gaps. Um, you know, it's like, you're imagining that you're looking at a photograph, you're like, Oh, just I want to look around the corner! I just want to see what's over there. What is it!

**Sara Mannheimer** 26:44

So then it's more just about scoping. It's like, this is what I have. And these are the connections I can make. And then you just explain that in your paper as you or in your lit review as you went through?

**QR04** 26:55

Yes. Except that, you know, not all of the people are dead. And so a thing that I also did was to sort of, you know, like, track people down and interrogate them very uncomfortably at conferences and stuff.... So listen, I've been reading all your stuff from 1973. And you mentioned what the [livestock] did. So where were they?

**Sara Mannheimer** 27:19

Did they remember?

**QR04** 27:21

Yeah, they usually did!

**Sara Mannheimer** 27:24

Okay, that's really interesting. Is that common practice in your field? Or?

**QR04** 27:30

I don't think so. Especially because people have a tendency to work within [one specific culture] and not cross culturally. So you know, I'll be like popping up and be like, so you're in X country in 1970. You know, what I mean, or whatever year. And and I need you to tell me what happened with this [resource]. People are surprised. Like, Oh, okay. You know? Or like, last week, I was reading your dissertation... and like gasp of horror. You know? And in Chapter Three, you didn't go very deep and describing this [resource] situation. Tell me more.

**Sara Mannheimer** 28:09

Hmm, that's very interesting. I mean, yeah, all of the data that we have is like, only partial, you know, so I'm not surprised that that would be something you'd have to do. But but that's what data curators...

**QR04** 28:27

But people do... They never think to read it. But they always notice because it's so important.

**Sara Mannheimer** 28:32

Yeah. Yeah, yeah, I feel like this is part of what data curators are trying to do. I mean, it wouldn't solve the problem of them not writing about the [resource] in a different research project, but it would solve problems where like, anything you would need to clarify with the original researcher, if they have died, like we want, we want there to be rich information to help. Okay, let's move to informed consent. So we talked a little bit about consent with, like Indigenous people and other people who are involved in the data. Can you tell me like a specific example of when you considered the idea of consent of the people who are documented in the [ethnographic data archive], and then what you did about it? If there were any strategies you use to sort of like, ameliorate your challenges?

**QR04** 29:29

Yes. So I would definitely say that, you know, my earliest interactions with the [ethnographic data archive] were maybe in the [late 20th Century]. And that was definitely before I had ever heard of Indigenous data sovereignty in any way. So, you know, my introduction into how to work with these ethnographic data predated, um, you know, my awareness of this idea of Indigenous data sovereignty. So I had definitely... by the time I understood what what was going on, I had definitely read things I think I was not supposed to have read. And I think I may have alluded to things, you know, although this wasn't the topic of my research, but I may have alluded to things where people did not provide proper consents. Um, you know, in particular, one of the things I study is [topic redacted], and so there are some pre 1950 ethnographies in the American West with American Indian groups that document this. And so I'm just I'm thinking no, that they don't have proper consent, right. And they were, you know, maybe like alluded to, in two words on it, the papers I'd written. And so I remember interacting with colleagues around this, and you know, having a conversation and being like, okay, that's not something I'm ever going to do again. And from now on, I only go to the published literature that's written by American Indian colleagues to find this kind of... right? So that was a shift that happened in my own scholarly practices, where it's like, now I'll go very deep into the published, you know, American Indian and Indigenous Studies journals, but I don't touch that stuff in the archives. Yeah.

**Sara Mannheimer** 31:17

Interesting. So you just make that choice as yourself. But anyone else could go in there. And like, Is there any, again, like any way that the archive? I don't know, like,

**QR04** 31:31

flags it?

**Sara Mannheimer** 31:32

Yeah. Or, you know, does the archive include the original consent form, so that you can review like what consent was given or anything like that?

**QR04** 31:41

Not the last time that I was in there, but you know, the the conversations I've had with my own American Indian colleagues, is that this stuff is everywhere, right? Like, it's in libraries, it's in newspapers, it's not just in the ethnographic record. And so for tribes that are implementing new data sovereignty procedures, you know, it's just a huge undertaking to find all this stuff where sacred information has been released, much less control it right. Yeah. So, um, I think it's a serious logistical problem. But I also think that, you know, for me as a methodologist, as a teacher, I can support those efforts, right. So we don't have to grab every archive, if we, you know, we hit the students, and we're like, Listen, you know, this stuff wasn't properly consented. people today who are working in digital data sovereignty don't want it written on. So we need to not do this if we don't have proper consent, even for archival data, right? Yeah. And so that's what I do in my own teaching. Like [in a national anthropology methods program has] a module on Indigenous data sovereignty. So everybody who comes through [that program] gets exposed to that, right?

**Sara Mannheimer** 32:54

How about for others, just like, other not as vulnerable populations that these ethnographies write about? How do you sort of, how do you think about consent for those people? Do you just assume that they were working with the researcher and that researcher was using the strategies that they were using? And yeah, what do you do?

**QR04** 33:16

Yeah, that's a really good question. So in my own work, I've always thought about my own ethical commitments, and what sorts of information am I willing to be a conduit for? Right? And so there are certainly things that I found in archives that I'm unwilling to be a conduit for.

**Sara Mannheimer** 33:35

Right.

**QR04** 33:36

But so I think what I'm trying to say, is clearly people were not consented to the standards that we say, and there's no going back, and there's no changing that, right. So some options are to not use the data at all. Right. And this is also true for the data that we're collecting today. We know that ideas of consent are constantly evolving, and probably the standards we're hearing today will make our data problematic in the future. Right. And like what we're doing with social media data right now, how is what's that going to look like in the future? Um, so, you know, I think we can interrogate our own ethical commitments today. And think about, you know, are we working in a way that advocates for the interests of people in these populations? Are we aligning with their needs, right? Are we being sensitive to express sensitivities? But yeah, an argument can also be made that that maybe we shouldn't be using those data, right?

**Sara Mannheimer** 34:46

Yeah. What do you... have you ever like reached out to certain communities whose data you're considering using like to put together like a little focus group or some kind of community advisory group to help you like... so that you don't have to make assumptions about what wouldn't would not be okay about to use?

**QR04** 35:04

Yeah. That's a really good question. So the way that I've used these archives, you know, there are hundreds of... at least hundreds of communities. And so logistically, No, I've never considered doing that. However, in our own research groups that I lead, we try to be very forward thinking and how we address community based participatory work and consenting. And we increasingly try not to publish at all unless we have members of the community as co authors. We create lasting relationships with scholars in the communities, so we don't have to exercise our own judgments about what would be acceptable to the communities. And so, and obviously, decisions about you know, what gets archived and where it's archived, those are conversations we have with members of the communities. So it does seem like the next step could be these conversations about Okay, things that have already been archived for how we feel about using those. Yeah,

**Sara Mannheimer** 36:08

I mean, yeah...

**QR04** 36:09

I think we... do have the infrastructure to at least begins have conversations about that.

**Sara Mannheimer** 36:16

That's something that's been talked about in the social media research community, too, is like, if you're pulling research data from certain communities, obviously, consent isn't possible in a big data context. So or I guess not, obviously, but in many cases, it's not. And so that's one of the ideas. So I was just wondering if it was happening in qualitative data too, and it seems like it is.

**QR04** 36:41

And it's not, I just want to say that obviously, these are not homogenized communities, we understand that they have very different and often needs and desires that are in conflict. Right. Yeah. But having these relationships in, particularly with scholars in the community to understand, you know, different perspectives and needs, helps us navigate those for sure.

**Sara Mannheimer** 37:01

Yeah. Yeah. Yeah. Okay. Great. Awesome. Let's move to question five out of seven, privacy and confidentiality. Can you tell me about a time, if any, during the this example, when you considered issues of privacy and confidentiality?

**QR04** 37:18

So just drawing from memory, and you know, this isn't the freshest memory because I think the paper that analysis, but the paper I'm talking about was published, I think, in [several years ago], but what I was looking for in these data, and you know, what I was pulling was largely cultural descriptions. And so they were rarely personal. Okay. Um, there are a few cases in the ethnographic record, where an entire ethnography is around one person, right? It's all this person's experiences. And then in that case, obviously, there's no privacy or confidentiality, it's all about this person. Um, yeah, but one presumes one hopes, that, that there was an appropriate relationship there. And that I mean, it is true that I've been telling you that there is quite a lot of problematic work in the early ethnographic record, there are also anthropologists who did very forward thinking work that was co published with members of communities, you know, right from the earliest scholarship, so you have models of, of work that I think would be considered, you know, quite ethical and cutting edge even today. So not all of it is problematic.

**Sara Mannheimer** 38:42

Yeah. I guess for such a big archive. Some of the questions, I'm like, thinking about are: was there deidentification, like as you went through these cultural records, like, did you see people named or, you know, did you think about, like, the size of the community? I guess, I'm just trying to think the size of the community and whether re identification would be possible or anything like that, and does do you know, if the [ethnographic data archive] has stuff like that.

**QR04** 39:16

I mean, so that... a lot of the early scholarship, you know, people were doing very detailed kinship charts, trying to understand, you know, how kinship guides naming conventions. So, there was a lot of I mean, the purpose of this scholarship was was identifying these kinship systems and naming conventions as they are definitely identifiable. also say that, you know, there is a reasonably well developed tendency in cultural anthropology for people to go back and do re studies of communities 20 years later, without the involvement of the initial scholar. And, and oftentimes, you know, re-analyzing and taking very strong issue with the initial findings. So that is an example of re identification.

**Sara Mannheimer** 40:07

Hmm, interesting.

**QR04** 40:10

You know, so what, you know, when a very famous case is Margaret Mead's work was really studied and you know the next person, you know, tried to essentially take apart her reputation as a scholar, that kind of takedown is very common.

**Sara Mannheimer** 40:26

Oh, I mean, I guess it's good to continue to study these populations, especially if, like common practice and ethical practice are changing over time, but I don't to like to hear about actual takedowns.

**QR04** 40:45

Yes, but you know, the other thing about ethnographers is that I think, um, historical models of ethnography, and I don't know the extent to which these are fully adhered to today for various reasons. But the idea is kind of like, or at least, the way I was trained, is that you're making a lifelong commitment to people in the communities with which you work. And it's very common to have what we call fictive kinship relationships. Right? So you have godchildren, you're contributing to the community financially over a period of decades. Right? So, you know, quite often you see critiques of the ethics of anthropologists, for example, like, Are people benefiting from from the work you did? Or you're getting a salary from this work? How are they benefiting? And so, historically, at least the way that I was trained is that, Yeah, there are supposed to be, you know, those remittances that are meant to last for the entirety of the career. And it's not like you go and you extract data, and you never come back. You're meant to go back every year, you know, like, until you die. Right. And so, my own mentors had these kinds of relationships, right. And I do in my own research in the sites where I work. So in terms of re identification, it's like, you could argue it's, you know, one 60-year long project.

**Sara Mannheimer** 42:09

Yeah. Right. And like, if the people in the communities where you work, like, trust you and talk with you, is there... Like, is there a sense that they want to be named? That they like, want their story to be told?

**QR04** 42:26

Yeah, a lot of times people want their story to be told. I mean, I know that in my work, I have those conversations people, right, like, do you want me to name your community? Or do you want me to give it another name? And if it's another name, what should we call it? Like, let's have a community meeting and discuss what we're going to call it? Or would you like to pick a name? Let's pick a name together, Right? So those are all conversations I have? Um, I wouldn't say that everyone does that. But I think a lot of people do.

**Sara Mannheimer** 42:53

Mm hmm. Very interesting.

**QR04** 43:01

But, I have not seen that in the metadata, just just gonna say, I haven't seen any metadata.

**Sara Mannheimer** 43:06

like deidentified or not deidentified...

**QR04** 43:08

you know, I mean, even in my own data, there's no place... I was never trained to document in a particular place what sorts of conversations... you know, you have your IRB forms, and so the IRB requires certain procedures, but the IRB does not ask about the identification of the community in publication. So that's how, and, you know, usually someone will put in their peer reviewed publications, you know, I've called the community this, and it's located in this general geographic region. So you can find it in the peer reviewed publications, but it's not necessarily a basic part of the data release.

**Sara Mannheimer** 43:42

Uh huh. And have you released your own qualitative data in addition to reusing?

**QR04** 43:50

So this is I wondered if you were going to ask me this. Um, you know, when I was trained, the IRB would make you promise to destroy the data. You know, the basic form was like this data destruction form, right. So—destroy it in five years destroy it in ten years. And I remember was like, I don't want to destroy it!

**Sara Mannheimer** 44:07

Yeah. Yeah, especially if you're working with the community your whole life!

**QR04** 44:11

They say, "you should never release it. It shouldn't even exist anymore." Right. And so it's only relatively recently that there's been this shift with the expectation that you might release the data. So in my own data practices that started with, you know, releasing very detailed codebooks that had typical exemplars pulled from the data, right, but not the actual, the data set. You know, now with large [federally] funded proposals, there is some expectation that you would consider it, but I think speaking for the really detailed qualitative data, no.

**Sara Mannheimer** 44:47

Now, yeah, yeah.

**QR04** 44:49

People are releasing their coded data, they're releasing their quantitative data, and much less so, you know, the transcriptions or the field notes. Ethnographers have a disciplinary history of leaving their ethnographic materials, everything, to libraries, right. And so that's where you see these practices.

**Sara Mannheimer** 45:10

I see.

**QR04** 45:11

Because they did a data release on a project, you know, because as I just explained, ethnographers don't usually think about their research, you know, in a [single] grant, they think of it as 30 years or 40 years. And so they'll take all that stuff and donate it to the library or university or something like that.

**Sara Mannheimer** 45:26

I see. More like a historical record. Yeah. Okay. All right. Our last theme, intellectual property. Did you think about the intellectual property either I guess from the the [ethnographic data archive], from their perspective, like the archive that holds them? Or for the people themselves who are part of the data? Did you think about intellectual property at all? And if so how?

**QR04** 45:55

So the [ethnographic data archive] does provide guidance that's flagged about how they want to be cited. Right. So that is there. Um, and, and so I adhere to those.

**Sara Mannheimer** 46:06

Mm hmm.

**QR04** 46:09

The other part about the intellectual property of the people who are in the studies, the second part of the question. I confess that I never thought about it that way until I started to learn about the, you know, the Indigenous data sovereignty literature. And then, it's like this total worldview shift, and it got me thinking about data in a very different way.

**Sara Mannheimer** 46:31

Yeah. Um, how about the researchers themselves? You know, would do you cite each researcher individually or just the [ethnographic data archive] as a whole?

**QR04** 46:42

So the way they have... you would cite the researcher, and then you would say, you know, that the data were published in the [ethnographic data archive].

**Sara Mannheimer** 46:50

Okay. I think that's enough, actually, for this question. So great. All right. Any additional issues or challenges that arose during your reanalysis of the data that I haven't asked you about that are applicable to data reuse and data curation?

**QR04** 47:11

Um, yeah, you know, the only other thing that I would say, is just that on some level, I think working between the [ethnographic data archive], and with colleagues who work on Indigenous data sovereignty made me really question, you know, on some level, um, a lot of the information in these ethnographies, is published in monographs, and they're in libraries, right. So it's not just what's in the in, you know, being held in archives, but also in libraries themselves, right. So there's been scholarly work that sort of cleared the bar of peer reviewed publication, and is also in this questionable space around consent and privacy, and ethics. And I haven't seen conversations about that at all. And, you know, I'm, this is not my area of scholarly expertise. I'm someone who avidly reads literature and I appreciate, you know, scholars who specialize in this area so that I can, you know, become more ethical in my own practices and teach my students better. Um, but yeah, I don't think that these problems end at the archives.

**Sara Mannheimer** 48:28

Yeah. Okay, great. Um, one final question. Sorry. Um, can you think of anyone else who I should interview other qualitative researchers who have reused data? Or who have published data? or social media researchers, if you know, of those?

**QR04** 48:53

So people I know, I mean, a lot of the people who I know who are working are computer scientists, right? They're working with social media, they, you know, we... So for us, as qualitative researchers, we always are looking at what computer scientists are doing, and trying to figure out, you know, how can we use these innovations? So there are some people I could you refer you to that work in that space?

**Sara Mannheimer** 49:16

That would be really helpful.

[Additional discussion of who else Sara should interview]

**Sara Mannheimer** 50:11

Yeah. Great. That would be really helpful. Thank you so much. This has been really great. It's just an it's so interesting to think about these issues from the different perspectives of different disciplines and everybody has a little bit of a different view on it. So I really appreciate your view.