Article

IFLA Journal
2026, Vol. 52(1) 48–58
© The Author(s) 2025
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/03400352251348439
journals.sagepub.com/home/ifl

Stewarding Indigenous language data: Case studies in CARE

Sandra Littletree
Information School, University of Washington, USA

Nestor Guerrero
Information School, University of Washington, USA

Miranda Belarde-Lewis
Information School, University of Washington, USA

Abstract

Stewarding qualitative Indigenous research data in libraries and repositories requires a nuanced, culturally responsive approach that respects Indigenous values and emphasizes relational accountability. This paper, using a case study approach, explores the decisions scholars of Indigenous language and culture face when depositing research data into a university-based special collections. The findings underscore issues of institutional trust and also indicate that providing access to their qualitative research data for future generations and community building is paramount. It also emphasizes the value of collaborative curation involving scholars, Indigenous communities, and the need for coordination between special collections and research data services within libraries. The aim of this research is to support libraries and repositories as they work to implement the CARE Principles for Indigenous Data for qualitative research data in alignment with Indigenous research methods and Indigenous knowledge systems.

Keywords

CARE, special collections, Indigenous Data, endangered heritage languages

Introduction

Following decades of disregard for Indigenous values and needs, the CARE Principles for Indigenous Data Governance (collective benefit, authority to control, responsibility, ethics) have provided guidance for the respectful stewarding of Indigenous research data, particularly in response to the open data movement in the quantitative data realms. The CARE Principles are meant to supplement the FAIR Principles (findable, accessible, interoperable, reusable) for scientific data management and stewardship, with the aim of offering enhanced guidance for researchers and data stewards of data collected in Indigenous contexts. However, more guidance is needed for information and data professionals to ethically steward qualitative data objects involving Indigenous cultures and languages, such as audio recordings, field notes, transcripts, and other descriptive metadata.

To put the CARE Principles into practice for qualitative data in research data services in academic libraries and repositories, the University of Washington Information School's Data Services for Indigenous Scholarship and Sovereignty (DSISS) project is working towards developing an Indigenous Data Services framework that will include customizable policies, protocols, and practices for university libraries to ethically steward Indigenous qualitative data. Our goal is to help move towards specificity by providing nuance to some of the lingering questions that abound when considering how to "Be FAIR and CARE."

Corresponding author:
Sandra Littletree, Information School, University of Washington, Box 352840, Mary Gates Hall, Suite 370, Seattle, WA 98195-2840, USA.
Email: Sandy505@uw.edu

As a component of the DSISS project, we present an analysis of two case studies of scholars of Indigenous language revitalization and culture who have collected qualitative data objects and who were bound by external funding requirements to deposit their data in an institutional special collection. Because of the potential cultural sensitivity of their data and the overall legacy of extractive and disrespectful research practices in academia, these scholars faced decisions about how and where they deposit their materials to maintain collective benefits that respect Indigenous communities and support their data sovereignty. Considering research methodologies that benefit Indigenous communities (Kovach, 2009; Smith, 2012; Wilson, 2008), the findings from these case studies indicate that providing access to their qualitative research data for future generations and community building is paramount. The findings reveal the value of a collaborative curation approach (Karcher et al., 2021) that engages scholars and Indigenous communities, adding complexity to the dimensions of collective benefit in CARE.

Background and literature review

Collecting Indigenous language and cultural material has been long considered a project of colonial pursuits, and the legacy of depositing these materials in libraries and special collections has created ongoing access and ownership issues for contemporary Indigenous communities. Librarians and archivists with Indigenous interests in mind have been working for decades to overcome the harms caused by colonial practices of mainstream information institutions through the creation of protocols or guidelines for engaging with Indigenous materials in collections (Aboriginal and Torres Strait Islander Data Archive, n.d.; Aboriginal and Torres Strait Islander Library, 2012; Callison, 2017; First Archivist Circle, 2007; International Council of Archives, 2019; Powell, 2014; Steering Committee on Canada's Archives, 2022). Additionally, the scholarship of Indigenous librarianship (Burns et al., 2010; Gosart, 2021) and archival practice, which emphasizes Indigenous well-being (Thorpe, 2022, 2024), relationality (Littletree et al., 2020), and decolonizing archives (Krebs, 2012; O'Neal, 2015), has worked in concert with guidelines to direct mainstream information institutions to handle Indigenous material with respect and to create points of access that are reflective of the world views and knowledge of Indigenous peoples in catalogs and finding aids.

The 2019 CARE Principles for Indigenous Data Governance (Research Data Alliance, 2019) and the International Indigenous Data Sovereignty movement have become a foundation for policy advocates to insist on the centering of Indigenous interests when researchers are collecting and disseminating data obtained in an Indigenous context. Information institutions, such as libraries and archives, are engaging with the CARE Principles as another tool to provide guidance for handling materials and data containing Indigenous knowledge (Chigwada and Ngulube, 2024; Cummins et al., 2023). The CARE Principles, which are focused on Indigenous data sovereignty goals, correspond with the FAIR Principles' emphasis on the open data movement. The phrase "Be FAIR and CARE" indicates the need for information professionals to create data management systems and policies that not only work for metadata standards and machine readability, but are also beneficial to Indigenous communities, especially when the data could be considered sensitive.

While commitment to Indigenous scholarship is increasing at research universities (Minthorn and Shotton, 2018), few have library and research data services designed to support the diverse practices of Indigenous scholarship. Wong (2024), writing about the overall lack of guidance focused on equity, diversity, and inclusion in research data management policies in academic libraries, notes the availability of the CARE Principles and the First Nations Principles of OCAP (ownership, control, access, possession; First Nations Information Governance Centre, n.d.) to help research data service librarians create more responsive research data management policies. However, the documented practice of upholding Indigenous data sovereignty goals for data repositories is lacking.

There are various challenges typically faced by archives and repositories that are responsible for stewarding language data. Scholars in the field of information science have identified deficiencies in authority control within language archive databases as being a prominent factor. These shortcomings in authority control are particularly evident in the proper names of people, where archives often maintain local name authority files without linking to widely recognized external authority systems (Burke and Oksana, 2020). Previous research has examined the role of librarians in Indigenous knowledge management (Sharief et al., 2021) and the challenges in digitized Indigenous knowledge management in the African context, including exposing sensitive data and the lack of metadata standards to provide access (Balogun, 2023). Previous research also indicates that anthropological qualitative research data held in archival settings poses discoverability problems for scholars as well as community researchers (Marsh, 2019).

Some linguists are moving forward without the use of institutional repositories to store and share their qualitative data, bypassing them altogether. These language warriors are adapting existing digital tools as they see fit—a phenomenon that coincides with the rise of various online and digital messaging, video and communication tools that are available immediately and at no cost to the user. As noted by Burke et al. (2022: 12), some linguists are "increasingly using social media platforms like YouTube, WhatsApp, and Facebook to disseminate materials to language communities due to their accessibility and ease of use, despite the known advantages of formal archives for sustainable preservation, security, and searchability." Their findings highlight a preference for tools that meet the needs of language communities. This refusal to use institutional data repositories and archives, which reflects a politics of refusal (Simpson, 2007), signals a need to open spaces of new possibilities that communities and researchers can trust.

Our previous studies (Belarde-Lewis et al., 2024; Palmer et al., 2022) reveal the complexity of operationalizing the CARE Principles, particularly when considering the contrasting and somewhat divergent goals of various stakeholders, including scholars, Indigenous communities, and institutions. In particular, the case of Salish language expert Tami Hohn (Puyallup) demonstrates the issues that arise when Indigenous language materials are held in a university special collection, without any protocols that allow the tribe to control access to their own language materials via the library catalog and finding aids (Palmer et al., 2022). The Hohn case provides the opportunity to reflect on how scholars and collecting institutions can enact responsibility and relational accountability through data and archival services that provide relevant, meaningful, and accessible materials for both academic and Indigenous communities.

Information professionals charged with stewarding qualitative Indigenous research data have much to gain from incorporating Indigenous perspectives into research data services. Scholars (Archibald, 2008; Cajete, 2000; Meyer, 2008; Wilson, 2008) have argued that Indigenous ways of knowing reflect the holistic nature of knowledge and the importance of relationships, community, place, and relational accountability, which can and should inform Indigenous data sovereignty practices (Duarte et al., 2020). Thorpe's (2022, 2024) examination of Indigenous archival sovereignty reveals the importance of relationships to place and people, as well as local protocols and governance structures. It is in this "space of cultural resurgence" that the archive can support "individual and community social, emotional and cultural needs" (Thorpe, 2022: 207). It is in this direction of acknowledging relational accountability in research data services that our project seeks to contribute. Drawing on the concept of contextual integrity (Nissenbaum, 2010), we present two cases to show the kinds of relationality associated with different methods and data products.

Methods

This study builds on previous findings of the University of Washington Information School's DSISS project. Since 2021, the research team has conducted iterative phases of stakeholder and participatory engagement, including a workshop with Indigenous scholars and researchers, local tribal experts, librarians (research data services, digital scholarship, special collections, ethnic studies), and experts in metadata management, data curation, and content management systems. Our previous analysis (Belarde-Lewis et al., 2024), based on engagement with scholars, Indigenous community members, and information and data professionals, explored themes of ownership, trust, and relational accountability. We indicated that the pursuit of CARE requires multifaceted and flexible data services, including operations within university special collections where Indigenous data is often held in archival collections. Building on our previous work, we again embrace the themes of ownership, trust, and relational accountability as a framework for our analysis of two cases of language researchers.

The cases presented in this article were selected based on the scholars' status as recipients of the Jacobs Research Fund. Originating in 1975 as a partnership between the Whatcom Museum and the University of Washington Libraries, the Jacobs Research Fund has supported more than 700 researchers studying Indigenous languages and cultures across the Americas, and consists of over 140 unique linguistic and anthropological field research collections derived from Indigenous/First Nations tribes of the Pacific Northwest, as well as regions in Canada, Mexico, and Latin America. One of the conditions of this grant is that recipients must deposit a final report and their data products (e.g. field notes, dissertations, audio recordings, spreadsheets) in the University of Washington Special Collections. Access to the materials is granted only with the donor's permission within 25 years of the date of the Jacobs Research Fund award.

We conducted an audit of the Jacobs Research Fund collection descriptions and metadata, and selected two scholars who are actively producing scholarship in the realm of Indigenous language revitalization and collecting born-digital research data. Strong

consideration was given to researchers whose field notes, description, and attribution metadata demonstrated a keen and intentional awareness of their relationality with the Indigenous peoples and communities involved in their project. The case studies feature an Indigenous researcher conducting projects within their own tribes/communities and a non-Indigenous Canadian researcher.

It is necessary to point out that the scholar in Case 2, although initially selected as a grant recipient, decided to decline the grant award because of their distrust of university archives in a non-tribal setting, resulting in their data sets not being deposited in the University of Washington Special Collections. We feature Case 2 because of the scholar's decision not to use the library as a data repository to contrast this with Case 1.

We held preliminary informational meetings with the scholars to provide context of our project's intentions. Following the informal meetings, we conducted recorded interviews discussing four key topics with the scholars: (1) the conceptualization behind their community-engaged project; (2) their research methodologies; (3) the outputs produced from their research data; and (4) their decision-making processes regarding data management and metadata standards, including the value and risks related to data storage and archiving, navigating expectations from multiple stakeholder groups (Indigenous communities, funding agencies, and collecting institutions/data repositories), and how they understand and are accounting for Indigenous data sovereignty within their research processes. The respondents gave their written consent for the interviews to be recorded and results to be published before the interviews began. The Human Subjects Division Ethics Review Committee at the University of Washington determined our research and interviews to be exempt (approval: STUDY00015479) on 22 April 2022.

The initial engagement and formal interviews with these scholars took place and were recorded using the Zoom videoconferencing platform, where we utilized its automated transcription feature to develop an initial draft transcript, which was then followed by manual revisions. These case studies were further expanded and analyzed through a collaborative process of transcription, writing, review, and revision involving the research team and the scholars.

Case study profiles

The following two case studies include researchers who were selected to receive the Jacobs Research Fund and worked on community-initiated research projects that produced qualitative data. The first case study is an ongoing institutionally supported project at a research university in Canada, and has culminated in a published dictionary and a web-based application that is downloadable to mobile phones. The second case was sustained by tribal community members in the USA, and has culminated in several modes of dissemination, including curricula and a published dictionary. The two cases are similar in the primacy of heritage Indigenous languages and differ in the approaches towards archival preservation. The profiles provide the background necessary for the preliminary analysis and discussion that follows.

Case 1: Coast Salish grammar research

The first case is a linguistic research project based in Canada and led by a non-Indigenous linguist. The original focus of this community-initiated language documentation project was to develop a grammar curriculum and eventually a dictionary for community use. The scholar's current research is studying the grammar of the language and lexical material, examining reduplication patterns, vowels, and semantics in verbs, and morphological changes. Pre-COVID-19, the bulk of the research was conducted during summer fieldwork; after the worldwide shutdown, the fieldwork shifted to regular meetings via videoconferencing.

The linguistic researcher and their team center collaboration with Indigenous community members throughout their research. The research and documentation required for a dictionary was at the request of the community. Additionally, the researchers have secured policy-based agreements in the research ethics review at their home institutions, and they share copies of all the data with the tribal officials. The research outputs include raw audio recordings of Native speakers, an online dictionary, and a database using Google spreadsheets that tracks metadata and transcription for audio files.

While the linguist and their team did not explicitly state that they used an archival framework based on digital preservation standards, their use of multi-site storage is a common archival and digital preservation practice (National Digital Stewardship Alliance, 2019). In addition to the copies stored in the clouddistributed to the tribal Nation, and held in the institutional data storage of the researcher, the project's data has been deposited in the University of Washington Special Collections.

Case 2: Language revitalization

The second case is based in the USA and showcases how the process of linguistic documentation

opportunities to record stories and conversations organically emerges. The researcher, as part of a four-person team, took a community-based approach to their collection of linguistic data as they were compiling a dictionary. The basis of their grant-funded project was focused on the language revitalization efforts of a local Indigenous language with the primary research outputs of a dictionary and associated curriculum. The researcher observed the outcomes of strengthened relationships between tribal members, as well as new opportunities for collective learning within the tribe, and noted an awareness that fortifying relationships are not considered "outputs" by western metrics of what it means to "produce" research. The researcher recognized these outcomes as significant and necessary to the team's process, and was aware of the positive and trust-building aspects of their research within their home community.

A key factor of the researcher's community-based research team was building a rapport with fluent language speakers throughout everyday and ceremonial interactions, prior to the more structured interviews. Instead of written or formal requests for interviews, a casual invitation was extended, usually involving a meal or relaxed setting in which to ask language-focused questions. Some of the team used what the researcher called "the sweat house method" of asking language and culture questions with groups of community members. A sweat house or sweat lodge can be a cleansing, meditative, and prayerful ceremony practiced by many Native American tribal communities. By participating in a sweat with community members, the team of "co-conspirators" built trust and demonstrated their commitment to the project to the community members who shared their expertise. The informal interviews were enhanced by structured and recorded interviews indoors with active note-taking. The researcher noted that this method and structure could have been utilized more often, but they made every effort to record their notes soon after their interviews. The team of community linguists worked for several years and published a dictionary of their tribal language.

Findings and analysis

We examine the differences and overlap in the scholars' approaches to navigating institutional trust, digital preservation, and relational accountability. Both cases are centered around language revitalization efforts through grant-funded research, supporting the global urgency that scholars and language workers approach their work knowing the consequences of the extinction of Indigenous languages (UNESCO, 2022). Both cases resulted in the creation of learning tools, such as a dictionary, a language application, or other type of online resource, to ensure the functionality and accessibility of their work. While both researchers in this study are concerned about how to get their research back into the hands of the community of origin, the cases represent two very different approaches to meeting this goal. Case 1 is highly structured, in an academic setting with copies of the language data being reinterred with institutional holding places as well as the governing body of the Nation. In contrast, the data and research outputs from Case 2 were not immediately deposited in an institutional repository or stored in a standardized format.

Institutional and personal trust

The two cases demonstrate different attitudes towards institutions as safe and trustworthy spaces for their research and research data. The scholar in Case 1 exhibits a sense of trust of the institution's stewardship of the research materials, including the permission process for granting access to the materials. The scholar states:

> My understanding or my assumption had been like, okay, either if it was a member of the Nations themselves, they [the University of Washington Special Collections] would either just give access or that they [the University of Washington Special Collections] would contact me, and then of course I would say yes. And, you know, help me help them find whatever it was. Look back through my own records. But I don't actually know. And it's never come up. And I think part of the reason it doesn't come up is that because I'm in such regular contact with everyone, if they're looking for something, they just ask me directly. I guess partly because I'm not exactly sure how all this works, and also it feels important to provide these materials very directly. We also have been giving the materials directly to the Nation. So they also have them. It's not only in the archive.

As this excerpt reveals, the scholar believes that enough trust has been built with the community that they know that community members can approach the researcher to obtain any data at any time. The scholar also trusts that the process of accessing materials through the University of Washington Special Collections would be relatively easy to navigate. Moreover, the scholar knows that the materials have also been deposited with the tribe, making multiple access points feasible for community members.

The Case 1 scholar also gives the qualitative language data from their research to the Nation's language coordinators. The scholar notes that there is not an official archive for the Nation, nor is there any librarian or information professional whose job it is to steward the data. This knowledge about the capacity of the Nation is part of the reason why this scholar advocated for having the data stored at the University of Washington Special Collections—"so the data can be stewarded long term." Here again is a signal of the trust the scholar has for the stewardship of their research data at the University of Washington Special Collections.

In contrast, the scholar from Case 2 expressed a strong distrust of institutions. The scholar had previous experience with other researchers who refused to share research data with the scholar and the tribal community, which developed into concerns about storing language and culture data with non-tribal institutions:

> Because of this, I am very, very hesitant to have any of my stuff or any of our stuff stored anywhere else. And frankly, and that's why I've been advocating for our own archives and a searchable database, because I just don't trust people and other institutions very well.

The Case 2 scholar's strong distrust resulted in their hesitation towards non-tribal institutions storing their qualitative research data. Eventually, their dissertation and other research data were archived at their local (tribally run) university, and they expressed an interest in backing up the data for future use.

Digital preservation

The scholars demonstrated differing approaches and challenges to the preservation of their born-digital data. In this article, we define "digital preservation" as the deployment of strategies and technologies to ensure that digital or digitized materials remain accessible and usable indefinitely (National Digital Stewardship Alliance, 2013). When asked about the storage of these massive and culturally important data sets, the Case 1 scholar ruminated on future access to the language data, rhetorically asking: "What will happen to all of this if/when I am no longer alive? Who will have access?" Despite this uncertainty, the scholar has worked to create multiple points of access for the research data, including sharing multiple hard drives with audio recordings with the tribal Nation and producing a cloud database (Google spreadsheets) storing metadata, in addition to depositing materials at the University of Washington Special Collections. The most common digital assets generated by the linguistics team represented in Case 1 are audio or video recordings, or a combination of both. When reflecting on the recordings, the Case 1 scholar stated:

> We ensure that these recordings are backed up on multiple hard drives and the Nation has access copies. Databases in the form of a cloud-based spreadsheet tool (Google Sheets) containing metadata and transcriptions for sound and video files (with contextual information of date and speaker) are created for the dictionary and grammar work being done. We use cloud tools to collaborate and not end up with conflicting versions. The Nations also have copies of these spreadsheets. This is how we keep track of what has been recorded and where we can look back at data we have documented.

The scholar in Case 2, on the other hand, employed what they referred to as "Salish pedagogy" for data preservation, reflecting traditional ways of passing down knowledge, where they described the capture of their preliminary data as "storing the notes and data in [their] head and heart." This approach prioritized intangible and organic methods of ethnography and documentation, emphasizing conversations and active presence with community members over the immediate act of recording notes during interactions. Instead, notes and preliminary data were created retrospectively from these engagements and contributed to the development of a Salish language dictionary. The Case 2 scholar used an ethnographic approach focused on protecting the intimacy and sacredness of their interactions with community members, sometimes at the expense of capturing a tangible recording, precise transcription, or active documentation. Nevertheless, they highlighted the importance of archiving data in a secure and sustainable repository, ideally built and managed by Indigenous communities. The research notes they recorded on paper were eventually stored on personal flash drives and as printed hard copies, which were shared between the scholar and their community collaborators. These early research products later informed an updated version of their Salish dictionary, which was ultimately stored locally in a tribal college database at the scholar's affiliated institution at the time.

The exchange below between the research team and the scholar in Case 2 exemplifies how the data was stored and their characterization of "Salish pedagogy" as a form of data preservation:

> Case 2 scholar: We have [the research notes/data], yes. We need to store them better, though. They're not all together … like, I have some, I know some of my co-conspirators—they have some … a bunch of it went into dictionary work. A bunch of it is just here [*points*

> *at head*] and here [*points at heart*] … like we just collected it, and we would share with each other and talk about it later. We would repeat it, things that we'd hear. So, it was kind of this constant conversation amongst ourselves and with the people we were relying on and sometimes they would ask us, "Hey, do you remember when I told you this?" … In a lot of Indigenous communities there's a lot of oral history, and people are asked to recite it or repeat it or bring it up … And so a bunch of it—we were just told stories … we were reminded and sometimes called to share some of this information at times.
>
> Interviewer: Because they knew that they shared it with you?
>
> Case 2 scholar: Yeah, yeah, yeah. It was sort of, like … I'll use the fancy word, the "Salish pedagogy" of "I've told you this story. Now you've got to tell others this story."

The scholar identified the risk of data loss due to the obsolescence of the born-digital media formats it was originally stored in as a significant challenge in digital preservation, driven by rapidly evolving technology. In the following excerpt of their interview, they describe how previous data (audio recordings, e-books, and curricula) did not migrate to newer software:

> I've had some experiences where, when I was at [tribal college] as a student, we experimented and we made some Salish language teaching materials on the old Macs and software programs that don't exist anymore. And now all I have are printouts of some of that stuff, like … you know, it didn't migrate, right? And also, there [were] some … really cool e-books that I worked on that have sounds. And I actually have recordings of elders telling the stories and some of them don't work now anymore. And that software is not supported anymore … So I've been thinking about this for a long time: How should stuff migrate? And how do we ensure that it migrates?

Hard-copy printouts were made of the e-book, but emulation could not be achieved. They described the digital resources they created for the project as having various interactive features like sounds and visual elements that readers could interact with—elements that were ultimately lost as the software it was created in was eventually discontinued, phased out, and relegated as obsolete. When describing previous efforts to document and digitize language materials, the researcher shared that the older materials were stored in outdated Apple software, and that much of their efforts and data was also lost. When considering the universal archival conundrum of ongoing preservation and the need to update their software, the researcher brought up the pertinent question of what robust and "fool-proof" migration strategies can look like in a world where digital technology continues to evolve.

Relational accountability

The two cases demonstrate the primacy of the scholars' positionality and relational accountability to the community of origin. Importantly, relational accountability goes beyond personal relationships. It is the responsibility for maintaining the well-being of the natural world, stories, objects, ancestors, and future generations (Wilson, 2008). As a form of relational accountability, the researchers engaged in research practices that maintained their responsibility to the community.

Both scholars highlighted the importance of returning the data to the community to ensure its relevance to the community's interests and needs. The Case 1 scholar, although not a member of the Indigenous community they centered in their research, emphasized the importance of the data for the generations to come, stating: "the audio recordings of Native speakers are the most invaluable resource for future generations." Similarly, the Case 2 scholar shared the following sentiment:

> And when I think about how, how do I want this knowledge out there? I want it to be held here by my community. You know. I want it to be in their mind, in their hearts. That's ultimately the most important thing.

These sentiments, reflective of the "collective benefit" component of CARE, demonstrate a desire to ensure that the language data becomes integrated into the generational knowledge of the community.

The scholar in Case 1 made decisions on when to record and when to stop recording when encountering potentially sensitive data, saying:

> There's always the chance that maybe some gossip got on there and I didn't realize. And I try to be really mindful of this and stop the recorder, you know, right when I anticipate this happening or see that it's happening.

The scholar in Case 2 demonstrated trust-building and developing relationality with their Indigenous community. In responding to a question about how they went about requesting consent from Native speakers in their community to document their knowledge via recording their conversations, they described the process of establishing a rapport with them and

communicating their intentions early on, planting seeds of trust and relationship-building over time:

> Interviewer: How did you let them [the Native speakers] know that you were recording, or that you were doing this in pursuit of research?
>
> Case 2 scholar: Well, we made our intentions known from years ago. There's three other primary people that I worked with for a long time … and then various others … We tried to stick to, like, what we know, like, a certain kind of cultural context. So, for example, there was a local kind of Salish-only dance. And we went there and we had made this announcement: "Hey, this is our plan. Our language is in jeopardy. We want to revitalize it. And we want any one of you to help us." And when we started a blanket dance—that was our first fundraiser, a blanket dance. We did a blanket dance and got some money, and that's how we first started. And then we would just go to people—speakers—and we would say: "Hey, this is what our intentions are … are you interested in helping us?"; "Can we interview you?"; "Can we ask you things?"; "Can we visit?" A lot of times we didn't say "Can we interview you?" Instead, we'd say "Hey, can we visit and learn from you?"

This excerpt clearly demonstrates the long-standing relationship between the researchers and the communities. The relational accountability the researchers have to the community, as well as to future generations, impacts their data collection processes. This kind of information is rarely included in data repositories to indicate the level of care researchers take when engaging with community knowledge.

Discussion

This study examined how two scholars of Indigenous languages and culture navigated institutional trust, digital preservation, and relational accountability. Although both were bound by the external funding requirements to deposit their data in the University of Washington Special Collections, one researcher (Case 2) ultimately made the decision to decline the Jacobs Research Fund award and keep the tribally based research outside of settler institutions such as libraries, special collections, and institutional repositories. The scholar's decision not to deposit and relinquish rights to their research materials and data (expressed in the deed of gift) was in itself a politics of refusal (Simpson, 2007)—a refusal of their data and research outputs to be stored indefinitely in the University of Washington Special Collections due to distrust resulting from historical harms that settler institutions have caused and perpetuated for Indigenous peoples; a refusal to operate through the usual detached objectivity and exhaustive documentation that is typical in colonial modalities of ethnography. Such instances of refusal open new spaces of possibility (Simpson, 2007) as they prompt a reshaping of our understanding of sovereignty and representation by valuing pauses, context, and respect for personal sovereignty over state-centered forms of recognition. As discussed in our previous work (Belarde Lewis et al., 2024), building institutional trust will take time, but it can only happen when Indigenous researchers and community members observe the sincere efforts of information and data scientists, librarians, archivists, and repository developers towards ethical stewardship. The CARE Principles —particularly the subprinciples of authority to control, responsibility, and ethics—provide a framework for responsible stewardship and trust-building.

The two cases in this article reveal the layered challenges and decisions that researchers may confront throughout the lifecycle of data. The importance of collaborative data curation becomes apparent as a means of enabling the ethically responsible reuse of data (Karcher et al., 2021). There is a need for collaborative alignment between scholars, Indigenous communities, and collecting institutions to develop tailored protocols for the long-term stewardship of qualitative cultural data (Belarde-Lewis et al., 2024). This kind of collaboration fosters dialog between archivists, data curators, researchers, and Indigenous communities to help identify the contextual gaps that must be addressed for the ethical stewardship of community-engaged research documenting Indigenous knowledge. The preservation of attribution, governance protocols, and the contextual accuracy of Indigenous knowledge within data repositories of partnering settler institutions occurs when relationships with Indigenous communities are maintained and sustained based on the needs of the community. This focus has emerged as a significant challenge for both collecting institutions and researchers. Collecting institutions often lack the necessary infrastructure or protocols to sustain relationships beyond the initial donor (the researcher). Meanwhile, scholars frequently do not establish clear protocols with their Indigenous partners on how governance over this data might endure beyond their involvement and be transferred to future generations of the community. Open dialogue, fostered through consultative and collaborative curation, helps these tensions surface and provides a pathway to addressing them.

The CARE Principles for Indigenous Data Governance include the principle of collective benefit, which stipulates that data ecosystems should be designed so that Indigenous people benefit from the data (Research Data Alliance, 2019). When attempting to operate

through the collective-benefit principle of CARE, researchers must carefully consider how their research outputs will be relevant, meaningful, and accessible to both academic audiences and the Indigenous communities they are engaging with and drawing knowledge from. It is imperative that researchers working with Indigenous data understand the nuances of this framework to ensure their approaches align with the CARE Principles.

In this rapidly evolving field of data governance, many questions remain about how western institutional archives and qualitative data repositories should restructure themselves to effectively adopt and implement the CARE Principles. Recently, emerging complementary frameworks to CARE, like SORT (situational awareness, outreach, repository protocols, technology), have been contributing to building out CARE from an aspirational framework towards operationalization by highlighting the distinct yet intertwining and complex roles, responsibilities, protocols, and processes that entail the stewarding of Indigenous data within a repository environment (O'Brien et al., 2024). This can be another step in offering guidance to institutions managing Indigenous data by specifying how these repositories can embody the indicators of the CARE Principles. However, the cases we have presented in this article demonstrate the nuances inherent in stewarding Indigenous knowledge in the form of qualitative language and culture data. They demonstrate that there is no one-size-fits-all approach, and these emerging frameworks are intended not as prescriptive solutions but as tools to articulate and develop nuanced protocols that respect the rights and interests of Indigenous peoples, even when there are differing interests. By examining some of the decisions that the scholars of Indigenous language and culture make about their data, information professionals working in research data services and special collections may begin to understand the nuances and challenges that emerge in the process of producing and preserving research data in alignment with the CARE Principles.

Conclusion

Our intention for this study was to deepen our understanding of how data services in libraries and repositories can best support Indigenous scholarship and knowledge sovereignty and fortify ongoing relationships of trust with Indigenous peoples. The cases reveal the complex decisions that scholars of Indigenous culture and language make as they collect and deposit their research data into repositories. Both scholars, highly aware of the need for preservation and access to their research data, also discuss the imperative of building strong relationships with the community. The analysis emphasizes the value of a collaborative curation approach that engages the needs of scholars and Indigenous communities in the data stewardship process. Communities need to be the ones to decide how repair should happen, not institutions.

The findings of the study gesture towards the importance of coordination between special collections and research data services within libraries to support the data sovereignty goals of Indigenous communities. Information professionals, including research data service librarians and special collections librarians, play an important role in promoting an ethic of care (Caswell and Cifor, 2016) and upholding the CARE Principles. Because these information professionals provide education and resources to researchers throughout the research data lifecycle, including consultation on data management protocols and data curation, they can be the first line of defense in upholding the CARE Principles for research data services and qualitative data repositories. Building on the work of Thorpe (2022, 2024), the DSISS project aims to move data repositories towards Indigenous well-being and Indigenous archival sovereignty. Research data services in academic libraries must have content specialists, resources, and the infrastructure to ethically support the culturally specific management of research projects involving Indigenous data.

Acknowledgments

We extend our gratitude to the scholars who shared their research practices, experiences, and perspectives. We thank our collaborators and partners at the University of Washington Libraries. We thank Dr Carole Palmer for her advisory contributions and Sadaf Bandeali for her editorial review.

Declaration of conflicting interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

This research was supported by a grant from the Andrew W. Mellon Foundation's Public Knowledge Program (grant number: G-2104-10533).

ORCID iDs

Sandra Littletree https://orcid.org/0000-0002-8291-5436
Nestor Guerrero https://orcid.org/0009-0003-8259-7989
Miranda Belarde-Lewis https://orcid.org/0009-0007-3064-9985

References

Aboriginal and Torres Strait Islander Data Archive (n.d.) ATSIDA protocols. Available at: http://www.atsida.edu.au/protocols/atsida (accessed 11 February 2025).

Aboriginal and Torres Strait Islander Library, Information and Resource Network (2012) Welcome to the ATSILIRN protocols for libraries, archives, and information services. Available at: https://atsilirn.aiatsis.gov.au/protocols.php (accessed 11 February 2025).

Archibald J (2008) *Indigenous Storywork: Educating the Heart, Mind, Body, and Spirit*. Vancouver, Canada: UBC Press.

Balogun T (2023) Data management of digitized Indigenous knowledge system in repositories. *Information Development* 39(3): 425–438.

Belarde-Lewis M, Littletree S, Braine I, et al. (2024) Centering relationality and CARE for stewardship of Indigenous research data. *Data Science Journal* 23: Article 32.

Burke M and Oksana Z (2020) Identifying challenges for information organization in language archives: Preliminary findings. Available at: https://digital.library.unt.edu/ark:/67531/metadc1755874/ (accessed 11 February 2025).

Burns K, Doyle A, Joseph G, et al. (2010) Indigenous librarianship. In: Bates MJ and Maack MN (eds) *Encyclopedia of Library and Information Sciences.* 3rd ed. Abingdon: Taylor and Francis, 2330–2346.

Cajete G (2000) *Native Science: Natural Laws of Interdependence*. Santa Fe, NM: Clear Light.

Callison C (2017) Truth and Reconciliation Committee report and recommendations. Available at: http://cfla-fcab.ca/wp-content/uploads/2017/04/Truth-and-Reconciliation-Committee-Report-and-Recommendations.pdf (accessed 11 February 2025).

Caswell M and Cifor M (2016) From human rights to feminist ethics: Radical empathy in archives. *Archivaria* 81: 23–43.

Chigwada J and Ngulube P (2024) Librarians' role in the preservation and dissemination of Indigenous knowledge. *IFLA Journal* 50(2): 242–256.

Cummins J, Soto A, Anderson J, et al. (2023) Enhancing stewardship of indigenous peoples' data, information, and knowledges in libraries and archives through Indigenous data governance. *Library Trends* 72(1): 13–47.

Duarte ME, Vigil-Hayes M, Littletree S, et al. (2020) "Of course, data can never fully represent reality": Assessing the relationship between Indigenous data and IK, TEK, and TK. *Human Biology* 91(3): 163–178.

First Archivist Circle (2007) Protocols for Native American archival materials. Available at: http://www2.nau.edu/libnap-p/protocols.html (accessed 11 February 2025).

First Nations Information Governance Centre (n.d.) The First Nations principles of OCAP®. Available at: https://fnigc.ca/ocap-training/ (accessed 11 February 2025).

Gosart U (2021) Indigenous librarianship: Theory, practices, and means of social action. *IFLA Journal* 47(3): 293–304.

International Council of Archives (2019) Tandanya—Adelaide Declaration. Available at: https://www.naa.gov.au/sites/default/files/2020-06/Tandanya-Adelaide-Declaration.pdf (accessed 11 February 2025).

Karcher S, Kirilova D, Pagé C, et al. (2021) How data curation enables epistemically responsible reuse of qualitative data. *Qualitative Report* 26(6): 1996–2010.

Kovach M (2009) *Indigenous Methodologies: Characteristics, Conversations and Contexts*. Toronto: University of Toronto Press.

Krebs AB (2012) Native America's twenty-first-century right to know. *Archival Science* 12(2): 173–190.

Littletree S, Belarde-Lewis M and Duarte M (2020) Centering relationality: A conceptual model to advance Indigenous knowledge organization practices. *Knowledge Organization* 47(5): 410–426.

Marsh DE (2019) Research-driven approaches to improving archival discovery. *IASSIST Quarterly* 43(2): 1–9.

Meyer MA (2008) Indigenous and authentic: Hawaiian epistemology and the triangulation of meaning. In: Denzin NK, Lincoln YS and Smith LT (eds) *Handbook of Critical and Indigenous Methodologies*. Los Angeles, CA: Sage, 217–232.

Minthorn RS and Shotton HJ (2018) *Reclaiming Indigenous Research in Higher Education*. New Brunswick, NJ: Rutgers University Press.

National Digital Stewardship Alliance (2013) Glossary. Available at: https://ndsa.org/glossary/ (accessed 11 February 2025).

National Digital Stewardship Alliance (2019) Levels of Digital Preservation. Available at: http://ndsa.org//publications/levels-of-digital-preservation/ (accessed 11 February 2025).

Nissenbaum H (2010) *Privacy in Context: Technology, Policy, and the Integrity of Social Life*. Stanford, CA: Stanford University Press.

O'Brien M, Duerr R, Taitingfong R, et al. (2024) Earth science data repositories: Implementing the CARE principles. *Data Science Journal* 23: Article 37.

O'Neal JR (2015) "The right to know": Decolonizing native American archives. *Journal of Western Archives* 6(1): Article 2.

Palmer CL, Belarde-Lewis M, Hohn T, et al. (2022) Indigenous research data case study: Toward contextual integrity for Indigenous data. Available at: http://hdl.handle.net/1773/49173

Powell TB (2014) The American Philosophical Society: Protocols for the treatment of Indigenous materials. *Proceedings of the American Philosophical Society* 158(4): 411–420.

Research Data Alliance (2019) CARE principles for Indigenous data governance. Available at: https://www.gida-global.org/care (accessed 11 February 2025).

Sharief OAE, Mudawi MSE and Mohamed RA (2021) Indigenous knowledge in Sudan: Perceptions among Sudanese librarians. *IFLA Journal* 47(3): 361–374.

Simpson A (2007) On ethnographic refusal: Indigeneity, "voice" and colonial citizenship. *Junctures* 9: 67–80.

Smith LT (2012) *Decolonizing Methodologies: Research and Indigenous Peoples*. 2nd ed. London: Zed Books.

Steering Committee on Canada's Archives (2022) Reconciliation Framework: The Response to the Report of the Truth and Reconciliation Commission Taskforce. Available at: https://archives2026.com/wp-content/uploads/2022/02/reconciliationframeworkreport_en.pdf (accessed 1 June 2025).

Thorpe K (2022) *Unclasping the White hand: Reclaiming and refiguring the archives to support Indigenous wellbeing and sovereignty*. PhD Thesis, Monash University, Australia.

Thorpe K (2024) Returning love to ancestors captured in the archives: Indigenous wellbeing, sovereignty and archival sovereignty. *Archival Science* 24(2): 125–142.

UNESCO (2022) Indigenous languages decade (2022–2032). Available at: https://www.unesco.org/en/decades/indigenous-languages (accessed 11 February 2025).

Wilson S (2008) *Research Is Ceremony: Indigenous Research Methods*. Halifax, NS: Fernwood.

Wong S (2024) Equity, diversity, and inclusion in institutional research data management strategies: A contributive justice approach. *International Journal of Librarianship* 8(4): 97–113.

Author biographies

Sandra Littletree (Navajo/Eastern Shoshone) is an assistant professor focused on Native North American Indigenous Knowledge at the Information School at the University of Washington. Her research focuses on the emerging field of Indigenous information science, particularly Indigenous librarianship and the intersections of tribal sovereignty, technology, knowledge, and information in Native North America. Her research is guided by Indigenous ways of knowing, that is, the ways Native people have been creating, transmitting, categorizing, and preserving knowledge since the beginning of time. Relationality is at the core of this approach, informing the structure, core values, and ethics of Indigenous information science. She examines institutions not just as repositories of information, but also as spaces that can maintain and support the continuation of Indigenous ways of knowing.

Nestor Guerrero (he/him) is a Queer Chicano archivist dedicated to sustaining libraries and archives that center memory work, local knowledge, accessibility, and creative practice. He currently serves as the project archivist for MARKINGS: Inscribing Indigenous Immigrant Oral Histories in Westlake/Pico-Union, a Mellon-funded initiative led by Heart of Los Angeles (HOLA), UCLA cityLAB, and local Mayan and Oaxaca community organizations. His work focuses on digital asset management and the stewardship of community-generated metadata. His experience in archives spans across academic libraries, arts and culture nonprofits, and community-powered radio. He has his MLIS from the University of Washington Information School (2024) and BA, Chicana/o & Central American Studies, from UCLA.

Miranda Belarde-Lewis (Zuni/Tlingit) is an associate professor and the inaugural Jill and Joe McKinstry Endowed Faculty Fellow of Native North American Indigenous Knowledge at the University of Washington's Information School. Her research focuses on the ways knowledge is documented and transmitted through Native art, and how information institutions are working to protect and respond to Indigenous data concerns. She is an independent exhibition curator who works with tribal, state, federal and international institutions and organizations to promote Native artists and their work. Her writing has been extensively published in numerous exhibition catalogs, art publications, and scholarly journals.