Managing Data, Managing Contradictions: Archiving and Sharing Ethnographic Data

---

Wolfgang Kraus and Igor Eberhard

During the last two decades, the Open Science movement has swept across academia, with massive repercussions for research and publishing. Starting in the natural sciences and engineering, it has brought wide-ranging promises of better science – more transparent, more reliable, more reproducible, more replicable, more efficient, more accountable, more relevant. A basic requirement to make all this possible, the argument goes, is the opening and managing of research data.

While the social sciences and humanities have been slower to embrace the Open Data paradigm – not least because the category of data itself is controversial – the increasing call for Research Data Management (RDM) and open access to research data has become a major concern in all scientific fields (see Allianz 2010), epistemological and methodological differences notwithstanding. The debate on researchers' responsibilities in producing, handling, and sharing their data, under such headings as Open Research Data or FAIR Data (FORCE11 n.d.), has also been taken up in Social and Cultural Anthropology and related disciplines that use qualitative methodologies such as ethnographic research (see, e.g., Imeri 2017; 2019; Mosconi et al. 2019; Pels 2018). In striking contrast to the optimism outlined above, researchers in these fields often perceive the call for Open Data as intrusive, enforced by research policies and funding agencies but difficult to reconcile with research practices based on relations of trust and, often, confidentiality.

Social and Cultural Anthropology is the discipline in which these reservations arguably are most acute, for several reasons. One is a history connected to European expansion and colonialism which has given rise to a heightened awareness of the ethical dimension of the relation between researchers and the people they study. Another is related to its defining methodology, ethnography, whose practitioners consider those being researched not as objects of study but rather as active collaborators in the construction of knowledge. As a major consequence of this, the understanding of research data in ethnography is very different from the one dominating much of the Open Data discourse (see Pels 2018).

Many ethnographers are deeply skeptical about handing over their data to others, and with good reasons (Imeri 2017; 2019, 49 f.). They fear that submitting to the increasing demand for Open Data may have a damaging impact because it risks undermining the social relations between researchers and research subjects that form the basis of ethnographic research. They also fear that it may compromise the ethical standards they aim to uphold in research and publishing. Giving away data means giving up control about what is done with them, while many researchers feel they can never relinquish the responsibility for the ways in which their data are used.

We argue that the skepticism of many ethnographers and qualitative researchers in the social sciences about data management and sharing is best understood as a result of a series of contradictions that have been discussed extensively in recent years but nevertheless remain underappreciated in much of the RDM policy discourse. A clear contradiction exists between Open Data and data protection (with the usual mitigating strategy being pseudonymization). However, the fact that ethnographic data are best when they are context-rich clashes with the ideal of avoiding personal identification, since pseudonymization only works if data are stripped of context information, or not at all. Furthermore, if ethnography is understood as a collaborative production of knowledge, it is contradictory that only one of the two sides involved – the researchers, but not the research subjects – should have control of the data and decide how they are to be managed.

To a large extent these and other contradictions discussed below remain underappreciated because, just like any field of social interaction, academia too is permeated by inequalities of power. One aspect of this fact is that some disciplinary areas have more power to define agendas such as Open Data, and the policies adopted in their interest: natural sciences more than the social sciences and humanities, quantitative methodologies more than qualitative, and so on. To put it bluntly: ranking low in the hierarchy of disciplines, the qualitative social sciences so far have not had much of a say in the Open Science debates. It is therefore of vital importance that their professional associations make themselves heard by formulating and communicating clear positions in the debate on data management and sharing (see, e.g., DGSKA 2019; DGS 2019; EASA n.d.; see also Boog et al. 2018). Furthermore, researchers should discuss and clarify their principles, practices and experiences concerning these issues. The present chapter is meant as a contribution to this rapidly growing body of literature.

The position we take is that, while indeed much of the skepticism is well founded, there is a lot to be gained from preserving and sharing ethnographic research data, if for different reasons and with different priorities than the ones usually invoked in the Open Data discourse. Based on this conviction, we started in 2017 to set up a digital data archive at the University of Vienna, the Ethnographic Data Archive (*eda*), with the aim of developing strategies and competences for the long-term preservation of ethnographic data and for data management in ongoing research.[1] In this chapter we will comment on the issues raised above and discuss our guiding principles and the insights and experiences we gained in the process. While being a small initiative with limited work capacity, we feel that we are doing important, and in the German-speaking world even pioneering, work, but much of it is still provisional and remains in progress (Eberhard and Kraus 2018, 49 f.).

What are Research Data?

A presupposition of the RDM discourse is that all research is data-based. Clearly, this is a matter both of definition and epistemology. In the humanities the relevance of the research data concept is frequently called into question, at least the way it is understood in the natural sciences (e.g., Andorfer 2015; Drucker 2011; Hügi and Schneider 2013). Peter Pels (2018) similarly argues that in anthropology *research materials* only become *data* once they are commodified, and warns that "reductionist definitions of research data may erase the variability of scientific perspectives and research paradigms" (2018, 3).

An additional difficulty is that the term "data" often implicitly refers to digital data only (e.g., Schöch 2013; FU Berlin n.d.). In the most basic logic of research the format of information – say, of an interview recording – is irrelevant. Nevertheless, many of the possibilities and challenges of managing research data being discussed have to do with their digital format and what might be called, their computability. The problem with the narrow understanding of research data as digital is, however, that by definition all digital information consists of data. Therefore, the category of research data runs the risk of becoming meaningless except as a reference to format, and to a field of methodological practices and possibilities (e.g., the Digital Humanities).

This raises the question of how research data in general are defined. Given the diversity of disciplinary practices, RDM policy statements often resort to mere listings of forms of data being used in various fields (e.g., DFG 2015). Such an approach results in a circular argument when treated as a definition (e.g., Allianz n.d.): research data are defined by their role in the research process, while research itself is defined through the systematic use of data.

Real definitions, by making statements about what research data are supposed to be, tend to give more away than these uncommitted formulations. An often-quoted definition originating from the United States Office of Management and Budget reads: "Research data is defined as the *recorded factual material* commonly accepted in the scientific community as necessary to *validate* research findings …"

---

[1] See https://eda.univie.ac.at/.

(OMB 1999, our emphasis). A slightly expanded version has been picked up by several UK institutions (see, e.g., ESPRC n.d., itself quoted by other institutions). Another influential statement defines research data as "Data that are *descriptive* of the research object, or are the object itself" (University of Bath 2011, our emphasis).

We argue that such an understanding of research data is based on an insufficient model of the research process. On the epistemological level, this model assumes that the main characteristic of research data is their ability to document aspects of the real world in a factual or, at least, descriptive sense, independently of the specific research context that has produced them.[2] Although due consideration is sometimes given to disciplinary specificities and differences, the basic model is often that of the natural sciences and rests on data as quantifiable information. If data reflect the real world independently of their research context, then they are unproblematic to reuse in a different context.

Another assumption in the RDM and Open Data discourse combines epistemology with accountability: access to research data may serve "to reproduce and verify the results" of research, as Austria's main funding institution for basic research states in the context of its Open Access Policy (FWF n.d.). The third and related element is the economic rationale: once public money is invested to fund research, the results must be made openly accessible. "Taxpayer-funded research" is an important buzzword here, and one that raises interesting questions concerning the role of national boundaries with regard to access to knowledge (see, e.g., the US Alliance for Taxpayer Access,[3] or the Foreword in UKRI 2016, 2 for a British example). By the same logic, the data collected during research become assets that cannot be owned by researchers. Rather, they must be shared and the public (including other researchers) has a right to access and re-use them.

Taken together, these assumptions imply two sets of ideas that sit rather uncomfortably with the practice and self-understanding of ethnographic research: first, ideals of objectivity and replicability of research, as well as a sharp discontinuity between everyday knowledge/experience and research-based knowledge, both of which we consider epistemologically mistaken; and second, principles of cost efficiency and accountability about whose neoliberal thrust we have serious reservations. Moreover, we argue that ethnographic research concerns a significant category of others, the research subjects, who are absent from this model of research except when being conceptualized in a paternalistic manner as those whose privacy and rights must be protected.

---

[2] It is true that the FAIR principles include "a requirement to openly and richly describe the context within which those data were generated," but this mainly serves "to enable evaluation of its [sic] utility" for secondary use (Mons et al. 2017, 52). We use the term context in a wider, social sense and invoke other arguments for its indispensability (see below).

[3] https://www.taxpayeraccess.org/.

Characteristics of Ethnography

In order to put the assumptions just outlined into perspective, let us consider the main characteristics and assumptions of ethnographic research. Ethnographic methods are now being employed across a wide range of disciplines and may mean vastly different things (e.g., Kazubowski-Houston and Magnat 2018). We are here primarily referring to the methodological approach that is mainstream in present-day anthropology, but much of our discussion applies to all forms of ethnography and is relevant for other primarily qualitative methods as well.

To avoid being misconstrued, we wish to make it clear that we are not arguing for a separate epistemology for ethnographic research. Rather, we contend that its example serves to highlight the shortcomings of a positivist model of research that underlies much of the RDM and Open Data discourse (Eberhard and Kraus 2018, 48). It does so by providing an extreme case of how research is always embedded in social relations, and therefore always has a context that needs to be taken into account – something that is increasingly being understood in other fields too. As a medical statistician remarks: "The very production of data is … always relevant to its interpretation" (Barrowman 2018).

In anthropology ethnographic research is typically done over extended periods of time in close contact, collaboration, and exchange with research subjects. It is based on communicative relations with those being researched and sees them as active participants rather than passive objects of study. This and the fact that ethnographic research often deals with personal lifeworlds implies important issues of trust and responsibility.

Ethnography is an open-ended methodology relying on a flexible combination of tools. Hence, ethnographers tend to produce varied and multiple forms and formats of data. Whether in analog or in digital formats, ethnographic data tend to be technically diverse. In order to make sense, the various kinds of data produced must be interpreted in relation to each other and to the overall research context and experience. With the establishment of new digital tools, media, and ways of disseminating and sharing ethnographic knowledge, this characteristic has come to be discussed as "multimodal anthropology" (e.g., Collins and Durington 2018) in recent years.

The predominant understanding and practice of ethnographic research as based on collaborative relations between researchers and the people they work with has several fundamental implications for the understanding of research data and for RDM:

(1) Data are not simply "found" or "collected." They are co-constructed in a process of dialog between researchers and research subjects. They do not merely document facts "out there" but are representations that contain the voices and perspectives of both sides involved.

(2) Therefore, ethnographic data cannot simply be owned by researchers (and even less by their institutions). They also belong to the research subjects and their communities, who may have their own interest in the data.

(3) There are no raw, uninterpreted data in ethnography. The dialogic process of making data is by necessity a process of interpretation.

(4) As products of social relations and dialog, ethnographic data are neither objective nor subjective. Both of these notions are predicated on a model of knowledge presuming a clear distinction between the observer and the observed object, a distinction that is neither meaningful nor possible in ethnographic research. Therefore, ethnographic data are never simply "descriptive of the research object" (University of Bath 2011); their primary referent is the ethnographic relation between researchers and research subjects.

(5) Both sides involved are embedded in specific social and cultural circumstances, bringing these into the ethnographic encounter. There is a gradual difference but no discontinuity between ethnographic knowledge and everyday knowledge and experience. Tearing down this positivist division also opens up the path to incorporating multiple and diverse ways of knowing, as in the debate on Indigenous methodologies (e.g., van Meijl 2019).

(6) Ethical considerations take precedence over considerations of efficiency in ethnographic research. The heightened ethical awareness is a consequence of the historical context in which the ethnographic methodology emerged and the responsibility that comes with its practice. However, the "primary ethical obligation shared by anthropologists … to do no harm" (AAA 2012) does not result from a potentially paternalistic and condescending attitude of researchers knowing what may be harmful for research participants. Instead, it is a concomitant of the fact that ethnographers require their research subjects to engage with them in the collaborative production of knowledge.[4]

Our notion that the common distinction between "raw" and "processed" or interpreted data makes no sense for ethnographic data[5] might appear to contradict the position Pels takes when stating: "Anthropologists should … insist on making an epistemological distinction between 'raw' and 'processed data'" (Pels 2018, 4). However, Pels uses these terms in a different and atypical sense. By "processing" he does not refer to the predominant meaning of making data usable for analysis but rather to the task of preparing them for reuse by others – in the sense of stripping

---

[4] It is instructive to compare the ethics declaration of the German Anthropological Association with its sociological counterpart. While the former (Hahn et al. 2008) stresses responsibility and reciprocity, the latter (DGS 2017) gives priority to objectivity.

[5] Also see Barrowman (2018) who argues that even in quantitative science there is no such thing as raw data.

them of information that might be critical or sensitive in the hands of secondary users.

We agree that in many cases "[e]xtensive processing of raw materials (beyond mere anonymization) becomes inevitable if others are to reuse them" (Pels 2018, 4). However, we do not agree with Pels' argument that sharing data as a matter of principle cannot be reconciled with the research subject's "rightful claims to knowledge shared with researchers" (Pels 2018, 4). It can, provided that research subjects are explicitly conceptualized as forming part of the audience for whose access and reuse data should be prepared and archived. Ideally, they should also be included in the process of selecting data objects for archiving, assigning metadata and meaning to them, and defining access regulations.

Towards an Ethnographic Archive

The idea of setting up an archive for ethnographic data at the University of Vienna first occurred to us when we realized that the Department of Social and Cultural Anthropology was heading towards a major generational transition, with several colleagues bound to retire over the next years. When reflecting on this development, it seemed to make sense to preserve the ethnographic material they had gathered during their research careers, involving them in the process while they were still available. These considerations did not come up in reaction to the increasing call for RDM and open access to research data. Instead, they grew out of our own research experiences and focused on existing data from earlier research projects, mostly in analog formats, that were to be digitally preserved and made accessible for reuse. The insight that it was also necessary to support ongoing research and provide data management expertise followed later from our experience of working with historical ethnographic materials and from our engagement with the Open Research Data debates. Our initiative thus is not representative of the "'top down' policy push" that Mosconi et al. identify as a characteristic of Open Science, but rather of what they refer to as the "collegial desire to share data" (Mosconi et al. 2019, 756).

It took several initiatives and a couple of years until the data archive we had in mind was put into practice, first as a collaborative two-year pilot project of the Vienna University Library and the anthropology department. After the successful pilot phase it was made permanent, albeit with still limited staff resources: Igor Eberhard in a part-time position as archive manager and Jasmin Hilbert as a student assistant, with Birgit Kramreither, the head of the Social and Cultural Anthropology Library, functioning as coordinator and Wolfgang Kraus as scientific leader.

While the main assumptions of (mainstream) ethnographic research as outlined above are probably uncontroversial for most anthropologists, our archival activities and the strategies we devised are based on additional assumptions. Our basic premise is that ethnographic data are of intrinsic interest beyond the primary research context, for two main reasons. First, they are complex and rich in ways which are hardly ever fully exploited in the original analysis. This is often a matter not only of complexity but also of sheer quantity (as both present authors have experienced in their own research). Second, since ethnographic data come out of encounters situated in time and space, they are historical by nature. As a consequence of transformation and change over time, they may become interesting and relevant in unforeseeable ways, not only for researchers and their scientific communities, but also for those being researched. Both reasons, we argue, provide good grounds for preserving ethnographic data and for making them accessible and reusable beyond the original research context, and these are entirely unrelated to the rationale of the Open Research Data discourse.

A further – and for qualitative social researchers perhaps obvious – assumption is that data are meaningless without context. From this widely shared conviction, radically different conclusions can be drawn. Hirschauer (2014), for instance, invokes the context-dependence of data such as interview statements to argue against the call for archiving and reusing qualitative social science data. We take the opposite position (even if we agree about the practical challenges involved, and that data archiving must be a matter of responsible choice rather than mandate). Contrary to Hirschauer, we maintain that it is possible to retain enough context in the archiving process to make data relevant for future uses, and have defined strategies to do so.

The notion of context is a key concept in the qualitative social sciences, and particularly so in present-day anthropology. As Dilley states, “stress on context in interpretation is one of [anthropology’s] distinguishing features; and it is relied upon as an indispensable part of anthropological method” (2002, 438). Especially after the discipline’s shift to midrange theorizing, contextualization has become an important part of anthropological explanation/understanding, based on the “view that context is generated and negotiated in the course of social interaction and exchange” (Dilley 2002, 439). This is not the place to attempt to clarify the general understanding of context in anthropology, except to note that it forms the indispensable background for a more narrow and tangible notion of context, that of research context, which is essential to our approach to archiving.

Ethnographic knowledge is embedded in social relations and in complex corporeal experience. Specific data objects cannot represent this context by themselves, but must be linked back to it in order to make sense and be interpretable. Moreover, ethnographic data tend to be diverse, and different kinds and formats of data must be interpreted in relation to each other and to the overall research context. A guiding principle for our archival activities is therefore that the link between specific data and their research context must be retained and documented as richly as possible. We will outline below how we are trying to achieve this.

The Ethnographic Data Archive (*eda*): Objectives and Strategic Considerations

The *eda* pilot project started in early 2017 and was made permanent in 2019.[6] Beyond creating and maintaining a digital archive of ethnographic data, *eda* aims to develop best-practice models for preserving ethnographic data for reuse. Our work addresses a wide range of challenges such as: (1) Defining archival and metadata strategies and standards adapted to the specificities of ethnographic research; (2) testing and defining best practice digitization workflows; (3) networking and exchange with other data management and archival initiatives in related fields;[7] and (4) identifying the ethical and legal issues involved and proposing solutions. While our current emphasis is on archiving historical ethnographic materials, mostly in analog formats, we aim to develop a comprehensive research data management strategy for anthropology and related fields in the medium to long term future.

One of our guiding principles is that collaboration with researchers makes more sense than the mere administration of legacies. This is confirmed by experiences we made working with materials existing in the departmental ethnographic collection, with often insufficient metadata and context information. More importantly, this conviction is based on our holistic understanding of ethnographic data objects as representing an interactive research process rather than separate aspects of an independent reality. Therefore, all data must be linked to the research settings and the researchers' biographies, thus enabling archive users to take this context into account. Researchers themselves are in a privileged position to accomplish the task of contextualizing and interlinking the data objects in a comprehensive manner. This also includes explication of theoretical and epistemological assumptions and the political context of research, something which is more obvious when dealing with historical data (e.g., material manifesting colonial involvement and/or racist assumptions), but is relevant with all kinds of data.

---

[6] The *eda* website (https://eda.univie.ac.at/) gives an overview of the team, activities and cooperations; the *eda* team can be contacted at: eda.ksa@univie.ac.at.

[7] Our national and international partners include CIRDIS/University of Vienna, Institute for Social Anthropology/Austrian Academy of Sciences, Phonogrammarchiv/Austrian Academy of Sciences, Department of Folk Music Research and Ethnomusicology/MDW, Fachinformationsdienst Sozial- und Kulturanthropologie, Qualiservice/University of Bremen.

Another guiding principle is the respect and support for legitimate interests of research subjects and source communities in the data, their protection from harm, and reciprocity. Here again, researchers are better positioned than archivists or future users to assess the interests and risks involved with specific data sets. As Pels notes, it is their "ethical duty to control how research materials 'go public'" (Pels 2018, 5). In a collaborative and dialogic understanding of ethnographic research, research subjects or their descendants and communities should be included in these decisions as far as possible (something we have not yet been able to accomplish).

On the practical level, we aim at sustainability through optimized workflows, the use of appropriate (open) file formats, standardized procedures and metadata, and ongoing quality control. Devising best practice digitization workflows requires a balancing of conflicting demands. The amount of work and cost involved and the required storage space should be kept low, while the technical quality of a digital copy should be such that it can be expected to be taken as an adequate representation of the analogue original even several decades from now. This requirement, together with a high level of technical autonomy (in the sense of not having to hand over materials to external service providers), is the guiding principle in our digitization strategy. Finally, in line with the considerations above, we leave the decision of what to archive to the researchers, while offering our advice (e.g., concerning ethical or legal issues) when being asked for it.

Data Objects

Eda's data objects are archived using PHAIDRA, the "repository for the permanent secure storage of digital assets at the University of Vienna."[8] As noted above, ethnographic data come in multiple formats. In our archiving activities the main focus is on text (e.g., field notes, transcripts and diaries), images (e.g., diapositives, black-and-white negatives and photo prints) and audio recordings (e.g., interviews, narrations, recitations, music). So far, most of the material has been in analog formats, but some has been digital too, sometimes in obsolete file formats that need to be converted to archival formats. Preferred file formats are PDF/A for text, TIFF for images and AIFF or FLAC for audio material; sometimes it makes sense to also archive the original files. We have not yet started working with film and video due

---

[8] See: https://phaidra.univie.ac.at/. Our close cooperation with the PHAIDRA team has proven to be highly productive and a most pleasant experience. Several people have substantially contributed to establishing and developing *eda*, most important among them: Maria Seissl, Susanne Blumesberger, Raman Ganguly, Rastislav Hudak and Claudia Feigl. We are deeply indebted to them for their ongoing support.

to the multiplicity of file formats, the need for data compression, and our lack of technical expertise in this field.[9]

We did extensive tests with various digitization approaches and defined best practices and workflows for various kinds of material. It should be understood that digitization often is not merely a technical procedure but includes a judgement about the relevant aspects of an object. When copying a faded photograph, for instance, we must decide if we are mainly interested in its current appearance – the result of an aging process – or the original information that can be restored by proper illumination and by digital editing, or both. In each case, the optimal digital copy or copies will be different. Once again, the researchers can help to make such decisions.

Our digitizing and archiving workflows are also based on considerations of the relation between analog objects as potential carriers of knowledge and their digital representations, and the transformation from one state to the other. The example of the faded image shows how a conception of what constitutes the data object in relation to the research context must guide the digital representation. Another example is a Compact Cassette containing several field recordings. Is the single isolated recording the object we are interested in, or is it the cassette as an entity representing a specific time span in the field? We have opted for the second as our predominant perspective. As a consequence, we devised a new object category for the PHAIDRA repository, the "container object," which allows us to retain context by retaining the integrity of the analog object.

The container object is a data object consisting of several files that represent the same analog object. In the case of the Compact Cassette just mentioned, the format of the container object makes it possible to retain the original connection between the recordings. In addition, a cassette often comes with more or less consistent metadata in the form of notes written on it or the cardboard insert in the box. In that case, the container object consists of several audio files and photos of the two cassette faces and the insert.[10] Other container objects, for instance, represent photographs from the ethnographic collection of the department with extensive captions on the verso.

---

[9] For those objects that have already been ingested, see https://phaidra.univie.ac.at/search#?owner=ethnograpp95.

[10] E.g., https://phaidra.univie.ac.at/o:953026.

Metadata

Descriptive information is rarely neutral. Seemingly innocent labels such as established names attached to collectivities or technical terms could carry problematic perspectives, ethnocentric assumptions, or obsolete theory. For these and related reasons we gave a lot of consideration to metadata-related questions. Our guiding principle in this process was a conception of research data as situated in the particulars of time and space and in the social encounter of fieldwork. The main consequences for our metadata approach are, first, attention to the research and social context of specific data objects, and second, the obligation to include the research subjects and source communities in our potential audience.

For metadata the PHAIDRA repository relies on a linked-data approach combining several established metadata standards, including Dublin Core, BIBFRAME, SKOS and others. We suggested several additional metadata fields that we considered useful, which were then mapped to existing categories. The PHAIDRA team proved extremely flexible and supportive in helping us to develop an *eda*-specific metadata scheme and submit form.[11] The submit form is relatively self-explanatory in order to enable researchers, after a short introduction, to upload their own data and metadata. This option is further facilitated by the possibility to create project-specific metadata templates. However, regardless of how it is done, providing rich metadata for their data objects requires a considerable effort and time investment by researchers.

Given the complexity of many data objects, we introduced a clear distinction within metadata between several object categories, some of which might coincide in the case of a given object. We refer to them as (1) born-digital object, (2) digital copy, (3) first-order analog object, and (4) second-order analog object.

What is meant with these categories can be illustrated with a musical recording. An audio file from a digital recorder is a born-digital object. A tape recording on a Compact Cassette, once digitized, is the digital copy of the audio content of a first-order analog object, the cassette. When the recording is considered a representation of a specific instrument being played, the instrument is the second-order analog object. Each of these categories, as far as they can be distinguished in a given object, requires its own metadata. An analog object in the second sense may also exist in the case of a born-digital object. The *eda* metadata scheme lets us follow this approach, even if a consistent terminology – not necessarily using our working terms, which could be improved – has not yet been implemented.

---

[11] For examples of objects with fairly complete metadata, see https://phaidra.univie.ac.at/o:1069269; https://phaidra.univie.ac.at/o:1048725.

A second innovation is the notion of "context data."[12] While access to objects in PHAIDRA can be restricted or blocked (an indispensable feature for *eda*), a basic given of metadata in PHAIDRA is that they are always public and open, without access restrictions or authorial responsibility and rights. Our interest in making the research context transparent and tangible turned out to be in potential conflict with this policy. Personal information that helps to contextualize data may not be fit for public access. Background information may be too extensive for metadata and require a research effort that needs to be credited, and so on.

What we refer to as context data provides a pragmatic workaround for these issues. Context data are data objects – often in text format (PDF/A) – providing information on other objects that is not suitable for metadata because it is too complex, needs being protected, or requires authorship and copyright because it is based on personal research, interpretation, and opinion. As separate data objects, context data can have access restrictions; they are referenced in the metadata of the objects they help to contextualize. A significant category of context data is information on research projects and ethnographers' biographies and research trajectories, e.g., the biographical interviews that Eberhard conducted with Elke Mader, now retired professor at the Vienna Department of Social and Cultural Anthropology.[13]

Pseudonymizing and Informed Consent

When sharing data, there is a clear contradiction between the interest in retaining as much context as possible in order to keep data meaningful, and the need to protect research subjects' privacy and identity (see, e.g., Cliggett 2016, 243–245). An ethnographer's main ethical responsibility is to protect research subjects from harm and to respect relations of trust and confidentiality established during research. Personal identifiability is in conflict with these obligations in many research contexts.

With regard to personal information only aggregate data can by anonymous in the strict sense. When aggregation is not intended or does not make sense, as in ethnography and other qualitative methodologies, only pseudonymizing is possible, and is routinely practiced. However, in typically small-scale ethnographic research settings it is not enough to suppress or replace names because just about any information can provide identification cues. Trying to remove all of this information means a radical thinning of data and risks making them useless (see Eberhard and Kraus 2018, 45 f.).

---

[12] See Eberhard 2020a, 173 f.

[13] See https://phaidra.univie.ac.at/search#?collection=o:1146526.

An established instrument to handle such conflicts is informed consent. This notion itself, however, needs to be adapted to the logic of ethnographic research and to specific research settings. Originating from and modelled on a biomedical research logic, it carries assumptions whose universality must be questioned, such as a deeply individualistic and ethnocentric conception of risk and consent. Even though granting consent is in most cases a personal process, the underlying decision-making is often strongly socially embedded. This can be observed not only in small Indigenous communities but also in nearby and familiar settings such as rural Austria (Kraus and Seiser 2022, 104 f.). Paying attention to these processes makes it clear that in a given context people may value social visibility in a specific community as much as or higher than the more abstract interest in anonymity. How to balance these contradictory values is often a difficult question, and one that can only be answered satisfactorily in dialog with research subjects.

A second assumption is that consent is given beforehand, once and for all, something which may be better suited to protect institutions (universities, funding agencies, repositories, etc.) than those being researched. If research and the social relations enabling it only unfold in the process, how can we give the information in advance to make informed consent possible at all? Hence the notion of "processual consent" (e.g., Rosenblatt 1995, 148 f.) which is however more demanding to implement and difficult to document than the established model, and grants research subjects rights (e.g., revoking consent) that they are often expected to sign away otherwise. If taken seriously, processual consent means that research subjects must also be included in the data management and archiving process as far as possible (see DGSKA 2019, 2).

When archiving existing data, a further problem arises: they often come from research contexts where documenting consent was not yet expected. In such cases researchers may help to disentangle the processual aspects of research and the expectations of research subjects concerning the use of data. Even where consent has been documented, there is a high chance that the online sharing of data was not considered as an option (Zeitlyn 2012, 471). There are no abstract general solutions for these problems; viable compromises must be found for each single case. They may include the necessity of restricting or postponing access to data analogous to archival closure periods. As far as possible, research subjects and/or their communities or descendants should be included in these decisions. Researchers themselves may also have an interest in protecting their privacy, with consequences for access to data.[14]

---

[14] On the issues discussed in this section, see, e.g., Cliggett 2016; Eberhard 2020a; 2020b; Eberhard and Kraus 2018; Imeri 2017; 2018; 2019; Lederman 2016; Zeitlyn 2012).

Perspectives for Further Development

In order to manage these and other challenges in archiving and sharing ethnographic data, solutions, instruments and perspectives still need to be worked out. At this point, we can only briefly mention three of the most pressing problem areas that must be addressed in the future development of *eda*.

One is the need for graded access to data for various categories of users. A considerable part of ethnographic data is not suitable for open access. In order to avoid the necessity of defensively precluding access to all sensitive material – an option that would defeat the short to mid-term purpose of the archive – several graded levels of access to data must be defined and managed. Models of how this can be done have already been developed (e.g., Imeri and Sterzer 2018; Sterzer et al. 2018) and implemented (e.g., at the Qualiservice research data center[15]). In *eda*'s case, this will have to be done at the level of the PHAIDRA repository which does not yet offer the necessary functionality. In addition, it will also require some amount of staff resources to manage access. But this will be an important and necessary step to increase the relevance of the archive.

The second problem concerns the need for standardization of metadata, and particularly the lack of controlled vocabularies, such as thesauri. This is now widely perceived as a pressing issue with major ethical implications, given that established terms, categories and names often stem from colonial situations, contain assumptions of cultural stability that fail to reflect dynamic processes of change and redefinition, or even carry racist and pejorative connotations. Activities to improve this situation are taking place in various contexts, e.g., in the project GND for Cultural Data (GND4C) (DNB 2019). We are in touch with some of these activities, e.g., a project currently under way at the Fachinformationsdienst Sozial- und Kulturanthropologie.[16] We are also planning a compact model project concerning a subcollection of annotated photographs from the departmental ethnographic collection made by Friedrich Dörbeck during the Hydrographic Expedition to Siberia (1902–12).[17] Taking this material as an example, we intend to evaluate metadata, descriptive terms, ethnonyms etc. and compare them with existing thesauri and technical terminology in collaboration with partners in the region.

This project will allow us to demonstrate existing issues in specific detail and suggest solutions. It will also serve as an exercise in trying to involve research subjects or their descendants, regional experts and others close to the field in the processes of describing and contextualizing data objects – an important aim for our future activities, as recommended by Corsín Jiménez (2018, 4) and others.

---

[15] See note 12.

[16] https://www.evifa.de/de/ueber-uns/fid-projekte/gemeinsame-normendatei-gnd.

[17] https://phaidra.univie.ac.at/detail/o:1165511.

As a logical extension of the ethnographic approach, the use of collaborative methods of knowledge production in order to make sense of objects and information items have increasingly been put into practice in recent years (for a museum-based example, see Scholz 2017a; 2017b). The question of re-contextualizing data and objects through collaboration and of what might be termed Indigenous metadata bring up a third problem area with even more far-reaching challenges than those just mentioned. It is related to the insight in anthropology and beyond that there is no pure, non-situated knowledge. All knowledge is socially and culturally positioned and serves specific interests. Positionalities and interests are diverse and embedded in relations of power and, often, injustice. Based on such a perspective, Indigenous communities have increasingly begun to claim control over knowledge relating to themselves. Much of this debate revolves around the notion of Indigenous Data Sovereignty (IDS) (Kukutai and Taylor 2016; Walter et al. 2021), i.e. "the right of Indigenous peoples to govern the collection, ownership, and application of data about Indigenous communities, peoples, lands, and resources" (Rainie et al. 2019, 301).

The IDS advocacy and activism movement recently proposed the "CARE Principles for Indigenous Data Governance" to complement the FAIR Principles (RDA IG 2019; Carroll et al. 2020). They aim to exploit the momentum of the Open Data movement to further IDS while simultaneously addressing some of its issues as seen from Indigenous perspectives. CARE stands for "Collective Benefit, Authority to Control, Responsibility, and Ethics" (RDA IG 2019). While Open Data agendas, including the FAIR Principles, typically argue in terms of abstract advantages for science and society in the singular, the CARE Principles stress interests of and benefits for specific communities, acknowledging the "power differentials and historical contexts" that are ignored in the Open Data discourse (RDA IG 2019).

With its focus on communities, collective benefit and diversity, the IDS movement provides a significant and important corrective to the individualist assumptions dominating much of the debate on Open Data and Research Data Management, e.g., its understanding of ethics. At the same time, however, the simplistic dichotomy of "Indigenous" and "mainstream" principles or values employed by some of its main proponents (e.g., Carroll et al. 2020) risks reproducing modernist dualisms, glossing over the diversity within and between Indigenous communities which may manifest itself in equally diverse interests.

Nevertheless, these and other forms of claims to control over knowledge by those being researched[18] are important and productive from an ethnographic perspective. The questions they raise are not only challenging but also enriching for

---

[18] E.g., the "Traditional Knowledge Labels" aiming to define "attribution, access, and use rights for Indigenous cultural heritage" according to community-specific understanding and values (https://localcontexts.org/), which cannot be discussed here for reasons of space.

initiatives such as *eda* when it comes to elaborating and defining the categories of data ownership, control, access, and licensing.

Conclusion

The main assumption underlying *eda*'s activities is that ethnographic data constitute historically situated representations of aspects of a world in flux. As such they do have a value beyond the primary research context and should be preserved and shared. However, the dialogic character of ethnography and the access ethnographers gain to personal life-worlds raise important issues around confidentiality, privacy, and reciprocity.

The management and archiving of ethnographic research data is a highly dynamic field with many challenges and contradictions. Such contradictions cannot be resolved on the level of principles. Nevertheless, it is possible to find pragmatic but ethically sound compromises that enable us to archive and share (with various forms of restrictions) ethnographic data without harming our research subjects. After all, the situation is not unlike the challenges that arise in the process of ethnographic writing and publishing – challenges for which we routinely find pragmatic single-case solutions.

However, these inherent tensions require additional measures that *eda* will have to deal with in the future, and before archiving more sensitive data. In order to strengthen the collaborative aspect of data management, we need models of, and experiences with, the integration of research subjects into all steps of the archiving process. Finally, in order to attain our goal of developing a comprehensive research data management strategy for anthropology and related fields, we must engage with current data practices in research. All of this is difficult to accomplish with the current limited staff. Nevertheless, as a small initiative and in a field where the questions still clearly outweigh the solutions, we feel that we have already achieved a lot.

Peter Pels suggests there is more than one "reason to consider social science data as indigenous or global heritage" (Pels 2018, 3). In this perspective the preservation and sharing of ethnographic data is an ethical obligation. In a similar stance, the American Anthropological Association lists as one of the main ethical principles for anthropologists to "protect and preserve [their] records" (AAA 2012). In a collaborative and dialogic understanding of ethnography, this logically entails the inclusion of research subjects and their communities in the process in order to safeguard data that comply with both FAIR and CARE Principles for the benefit of all interested parties.

References

AAA. 2012. “Principles of Professional Responsibility.” *American Anthropological Association.* Accessed October 30, 2021. http://ethics.americananthro.org/category/statement/.

Allianz der deutschen Wissenschaftsorganisationen. 2010. “Grundsätze zum Umgang mit Forschungsdaten.” *Gfz-Potsdam.de*. Accessed October 30, 2021. https://doi.org/10.2312/ALLIANZOA.019.

Allianz der deutschen Wissenschaftsorganisationen (n.d.). “Forschungsdaten.” *allianzinitiative.de.* Accessed October 30, 2021. https://www.allianzinitiative.de/archiv/forschungsdaten/.

Andorfer, Peter. 2015. “Forschungsdaten in den (digitalen) Geisteswissenschaften: Versuch einer Konkretisierung.” *DARIAH-DE Working Papers 14, Göttingen 2015.* Accessed October 30, 2021. http://nbn-resolving.de/urn:nbn:de:gbv:7-dariah-2015-7-2.

Barrowman, Nick. 2018. “Why Data Is Never Raw.” *The New Atlantis* 56. Accessed October 30, 2021. https://www.thenewatlantis.com/publications/why-data-is-never-raw.

Boog, Igor, Henrike Florusbosch, Zane Kripe, Tessa Minter, Peter Pels and Metje Postma. 2018. “Data Management for Anthropologists and Ethnographers: A Position Paper.” *Social Anthropology/Anthropologie Sociale* 26 (3): 7–10. Accessed October 30, 2021. https://doi.org/10.1111/1469-8676.12526.

Carroll, Stephanie Russo, Ibrahim Garba, Oscar L. Figueroa-Rodríguez, Jarita Holbrook, Raymond Lovett, Simeon Materechera, Mark Parsons, Kay Raseroka, Desi Rodriguez-Lonebear, Robyn Rowe, Rodrigo Sara, Jennifer D. Walker, Jane Anderson and Maui Hudson. 2020. “The CARE Principles for Indigenous Data Governance.” *Data Science Journal* 19 (1): 43. Accessed October 30, 2021. https://doi.org/10.5334/dsj-2020-043.

Cliggett, Lisa. 2016. “Preservation, Sharing, and Technological Challenges of Longitudinal Research in the Digital Age.” In *eFieldnotes: The Makings of Anthropology in the Digital World*, edited by Roger Sanjek and Susan W. Tratner, 231–50. Philadelphia: University of Pennsylvania Press. Accessed October 30, 2021. https://doi.org/10.9783/9780812292213-014.

Collins, Samuel G. and Matthew S. Durington. 2018. “Multimodality.” In *The International Encyclopedia of Anthropology*, edited by Hilary Callan. Wiley Online Library. Accessed October 30, 2021. https://doi.org/10.1002/9781118924396.wbiea2304.

Corsín Jiménez, Alberto. 2018. *A Data Governance Framework for Ethnography v. 1.0.* Spanish National Research Council, November 2018. Accessed October 30, 2021. http://hdl.handle.net/10261/172227.

DFG. 2015. “Deutsche Forschungsgemeinschaft: Leitlinien zum Umgang mit Forschungsdaten.” *Deutsche Fördergemeinschaft.* Accessed October 30, 2021. http s://www.dfg.de/download/pdf/foerderung/antragstellung/forschungsdaten/richtlinien_forschungsdaten.pdf.

DGS. 2017. "Ethik-Kodex der Deutschen Gesellschaft für Soziologie (DGS) und des Berufsverbandes Deutscher Soziologinnen und Soziologen (BDS)." *Deutsche Gesellschaft für Soziologie.* Accessed October 30, 2021. https://soziologie.de/dgs/ethik/ethik-kodex.

DGS. 2019. "Bereitstellung und Nachnutzung von Forschungsdaten in der Soziologie. Stellungnahme des Vorstands und Konzils der DGS." *Deutsche Gesellschaft für Soziologie.* Accessed October 30, 2021. https://soziologie.de/aktuell/news/bereitstellung-und-nachnutzung-von-forschungsdaten-in-der-soziologie.

DGSKA. 2019. "Position Paper on the Handling of Anthropological Research Data, German Anthropological Association (GAA)." *Deutsche Gesellschaft für Sozial- und Kulturanthropologie.* Accessed October 30, 2021. https://www.dgska.de/wp-content/uploads/2019/12/Positionspapier_Bearbeitet-fu%CC%88r-MV_EN_29.11.2019.pdf.

DNB. 2019. "GND für Kulturdaten (GND4C)." *Deutsche Nationalbibliothek.* Accessed October 30, 2021. https://www.dnb.de/DE/Professionell/ProjekteKooperationen/Projekte/GND4C/gnd4c.html.

Dilley, R. M. 2002. "The Problem of Context in Social and Cultural Anthropology." *Language & Communication* 22 (4): 437–456. Accessed October 30, 2021. https://doi.org/10.1016/S0271-5309(02)00019-8.

Drucker, Johanna. 2011. "Humanities Approaches to Graphical Display." *Digital Humanities Quarterly* 5 (1). Accessed October 30, 2021. http://www.digitalhumanities.org/dhq/vol/5/1/000091/000091.html.

EASA. "EASA's Statement on Data Governance in Ethnographic Projects." *European Association of Social Anthropologists*. Accessed October 30, 2021. https://www.easaonline.org/downloads/support/EASA%20statement%20on%20data%20governance.pdf.

Eberhard, Igor. 2020a. "Der Kontext bestimmt alles: Kontextdaten und Containerobjekte als Lösungsmöglichkeiten für den Umgang mit sozialwissenschaftlichen qualitativen Daten. Erfahrungen aus dem Pilotprojekt 'Ethnographische Datenarchivierung' an der Universitätsbibliothek Wien." *ABI Technik* 40 (2): 169–76. Accessed October 30, 2021. https://doi.org/10.1515/abitech-2020-2007.

Eberhard, Igor. 2020b. "Herausforderung ethnographische Daten: Erfahrungen und Ergebnisse aus dem Pilotprojekt Ethnographische Datenarchivierung an der Universität Wien." In *Künstliche Intelligenz in Bibliotheken: 34. Österreichischer Bibliothekartag Graz 2019*, edited by Christina Köstner-Pemsel, Elisabeth Stadler and Markus Stumpf, 261–76. Graz: Unipress Verlag. Accessed October 30, 2021. https://doi.org/10.25364/guv.2020.voebs15.20.

Eberhard, Igor and Wolfgang Kraus. 2018. “Der Elefant im Raum: Ethnographisches Forschungsdatenmanagement als Herausforderung für Repositorien.” *Mitteilungen der Vereinigung Österreichischer Bibliothekarinnen und Bibliothekare* 71 (1): 41–52. Accessed October 30, 2021. https://doi.org/10.31263/voebm.v71i1.2018.

EPSRC. “EPSRC Policy Framework on Research Data: Scope and Benefits.” *Engineering and Physical Sciences Research Council*. Accessed October 30, 2021. https://epsrc.ukri.org/about/standards/researchdata/scope/.

FORCE11. “Guiding Principles for Findable, Accessible, Interoperable and Re-usable Data Publishing Version B1.0.” Accessed October 30, 2021. *force11*. https://www.force11.org/fairprinciples.

FU Berlin. “Was sind Forschungsdaten?” *Freie Universität Berlin*. Accessed October 30, 2021. https://www.fu-berlin.de/sites/forschungsdatenmanagement/ueber-forschungsdaten/forschungsdaten/index.html.

FWF. “Open Access to Research Data.” *Der Wissenschaftsfonds*. Accessed October 30, 2021. https://www.fwf.ac.at/en/research-funding/open-access-policy/open-access-to-research-data/.

Hahn, Hans Peter, Annette Hornbacher and Michael Schönhuth. 2009. “‘Frankfurter Erklärung’ zur Ethik in der Ethnologie.” *Deutsche Gesellschaft für Völkerkunde*. Accessed October 30, 2021. https://www.dgska.de/wp-content/uploads/2016/07/DGV-Ethikerklaerung.pdf.

Hirschauer, Stefan. 2014. “Sinn im Archiv? Zum Verhältnis von Nutzen, Kosten und Risiken der Datenarchivierung.” *Soziologie* 43 (3): 300–312. Accessed October 30, 2021. https://publikationen.soziologie.de/index.php/soziologie/article/view/795.

Hügi, Jasmin and René Schneider. 2013. *Digitale Forschungsinfrastrukturen in den Geistes-und Geschichtswissenschaften*. Genève: Haute école de gestion de Genève. Accessed October 30, 2021. https://doc.rero.ch/record/31535.

Imeri, Sabine. 2017. “Open Data? Zum Umgang mit Forschungsdaten in den ethnologischen Fächern.” In *E-Science-Tage 2017: Forschungsdaten managen*, edited by Jonas Kratzke and Vincent Heuveline, 167–78. Heidelberg: heiBOOKS. Accessed October 30, 2021. http://doi.org/10.11588/heibooks.285.377.

Imeri, Sabine. 2018. “Archivierung und Verantwortung: Zum Stand der Debatte über den Umgang mit Forschungsdaten in den ethnologischen Fächern.” In *Working Paper Series des Rates für Sozial- und Wirtschaftsdaten: Archivierung und Zugang zu Qualitativen Daten* edited by Regina T. Riphahn, Gert G. Wagner und Heike Solga, 69–79. Working Paper 267. Berlin: Rat für Sozial- und Wirtschaftsdaten. Accessed October 30, 2021. https://doi.org/10.17620/02671.35.

Imeri, Sabine. 2019. “‘Open Data’ in den ethnologischen Fächern: Möglichkeiten und Grenzen eines Konzepts.” In *Forschungsdesign 4.0 – Datengenerierung und Wissenstransfer in interdisziplinärer Perspektive*, edited by Jens Klingner and Merve

Lühr, 45–59. Dresden: Institut für Sächsische Geschichte und Volkskunde. Accessed October 30, 2021. https://doi.org/10.25366/2019.07.

Imeri, Sabine and Wjatscheslaw Sterzer. 2018. "Schwierige Daten: Forschungsdatenmanagement qualitativer Daten am Beispiel der ethnologischen Fächer" *Präsentation für 107. Deutschen Bibliothekartag in Berlin 2018.* Accessed October 30, 2021. https://nbn-resolving.org/urn:nbn:de:0290-opus4-36849.

Kazubowski-Houston, Magdalena and Virginie Magnat, eds. 2018. *Special Issue: The Transdisciplinary Travels of Ethnography. Cultural Studies ↔ Critical Methodologies* 18 (6). Accessed October 30, 2021. https://journals.sagepub.com/toc/csca/18/6.

Kraus, Wolfgang and Gertraud Seiser. (2022). "Lokalgeschichte und Identität: Eine sozialanthropologische Lehrforschung als partizipatives Projekt mit Museum und Gemeinde." In *Geschichte vor Ort: Salzburger Museen und Sammlungen*, edited by Dagmar Bittricher, Andrea Dillinger and Martin Knoll. Salzburg: Edition Tandem.

Kukutai, Tahu and John Taylor, eds. 2016. *Indigenous Data Sovereignity: Toward an Agenda.* Canberra: Australian National University Press. Accessed October 30, 2021. http://doi.org/10.22459/CAEPR38.11.2016.

Lederman, Rena. 2016. "Archiving Fieldnotes? Placing 'Anthropological Records' Among Plural Digital Worlds." In *eFieldnotes: The Makings of Anthropology in the Digital World*, edited by Roger Sanjek and Susan W. Tratner, 251–71. Philadelphia: University of Pennsylvania Press. Accessed October 30, 2021. https://doi.org/10.9783/9780812292213-015.

Meijl, Toon van. 2019. "Doing Indigenous Epistemology: Internal Debates About Inside Knowledge in Māori Society." *Current Anthropology* 60 (2): 155–73. Accessed October 30, 2021. https://doi.org/10.1086/702538.

Mons, Barend, Cameron Neylon, Jan Velterop, Michel Dumontier, Luiz Olavo Bonino da Silva Santos and Mark D. Wilkinson. 2017. "Cloudy, Increasingly FAIR: Revisiting the FAIR Data Guiding Principles for the European Open Science Cloud." *Information Services & Use* 37 (1): 49–56. Accessed October 30, 2021. https://doi.org/10.3233/ISU-170824.

Mosconi, Gaia, Qinyu Li, Dave Randall, Helena Karasti, Peter Tolmie, Jana Barutzky, Matthias Korn and Volkmar Pipek. 2019. "Three Gaps in Opening Science." *Computer Supported Cooperative Work* 28 (3–4): 749–789. Accessed October 30, 2021. https://doi.org/10.1007/s10606-019-09354-z.

OMB. 1999. "Office of Management and Budget Circular A-110." *Whitehouse.gov.* Accessed April 04, 2022.
https://www.whitehouse.gov/wp-content/uploads/2017/11/Circular-110.pdf.

Pels, Peter. 2018. "Data Management in Anthropology: The Next Phase in Ethics Governance?" *Social Anthropology/Anthropologie Sociale* 26 (3): 1–6. Accessed October 30, 2021. Accessed October 30, 2021. https://doi.org/10.1111/1469-8676.125 26.

Scholz, Andrea. 2017a. "Sharing Knowledge as a Step toward an Epistemological Pluralization of the Museum." *Museum Worlds: Advances in Research* 5 (1): 133–148. Accessed October 30, 2021. https://doi.org/10.3167/armw.2017.050112.

Scholz, Andrea. 2017b. "'Wissen teilen' als postkoloniale Museumspraxis – Ein Kooperationsprojekt zwischen der Universidad Nacional Experimental Indígena del Tauca (Venezuela) und dem Ethnologischen Museum Berlin." *Sociologus* 67 (1): 59–82. Accessed October 30, 2021. https://doi.org/10.3790/soc.67.1.59.

Schöch, Christof. 2013. "Big? Smart? Clean? Messy? Data in the Humanities." *Journal of Digital Humanities* 2 (3). Accessed October 30, 2021. http://journalofdigitalhumanities.org/2-3/big-smart-clean-messy-data-in-the-humanities/#big-smart-clean-messy-data-in-the-humanities-n-6.

Sterzer, Wjatscheslaw, Sabine Imeri and Matthias Harbeck. 2018. "Zugriff auf ethnologische Forschungsdaten: Anforderungen und Lösungen" *Poster für 107. Deutscher Bibliothekartag in Berlin 2018*. Accessed October 30, 2021. https://nbn-resolving.org/urn:nbn:de:0290-opus4-157193.

Rainie, Stephanie Carroll, Tahu Kukutai, Maggie Walter, Oscar Luis Figueroa-Rodriguez, Jennifer Walker and Per Axelsson. 2019. "Indigenous Data Sovereignty." In *The State of Open Data: Histories and Horizons*, edited by Tim Davies, Stephen B. Walker, Mor Rubinstein and Fernando Perini, 300–319. Cape Town: African Minds and Ottawa: International Development Research Centre.

RDA IG. 2019. Research Data Alliance International Indigenous Data Sovereignty Interest Group. "CARE Principles for Indigenous Data Governance." *The Global Indigenous Data Alliance*. Accessed October 30, 2021. https://www.gida-global.org/care.

Rosenblatt, Paul C. 1995. "Ethics of Qualitative Interviewing with Grieving Families." *Death Studies* 19 (2): 139–55.

UKRI. 2016. "Concordat on Open Research Data." *UK Research and Innovation*. Accessed October 30, 2021. https://www.ukri.org/wp-content/uploads/2020/10/UKRI-020920-ConcordatonOpenResearchData.pdf.

University of Bath. 2011. "ERIM Terminology V4." *Wiki.bath.ac.uk*. Accessed October 30, 2021. https://wiki.bath.ac.uk/display/ERIMterminology/ERIM%20Terminology%20V4.

Walter, Maggie, Tahu Kukutai, Stephanie Russo Carroll and Desi Rodriguez-Lonebear, eds. 2021. *Indigenous Data Sovereignity and Policy*. Abingdon, Oxon: Routledge. Accessed October 30, 2021. https://library.oapen.org/handle/20.500.12657/42782.

Zeitlyn, David. 2012. "Anthropology in and of the Archives: Possible Futures and Contingent Pasts. Archives as Anthropological Surrogates." *Annual Review of Anthropology* 41: 461–80. Accessed October 30, 2021. https://doi.org/10.1146/annurev-anthro-092611-145721.