

University of Kentucky

UKnowledge

Anthropology Faculty Publications

Anthropology

6-17-2013

Qualitative Data Archiving in the Digital Age: Strategies for Data Preservation and Sharing

Lisa Cliggett

University of Kentucky, lisa.cliggett@uky.edu

Follow this and additional works at: https://uknowledge.uky.edu/anthro_facpub



Part of the [Anthropology Commons](#)

[Right click to open a feedback form in a new tab to let us know how this document benefits you.](#)

Repository Citation

Cliggett, Lisa, "Qualitative Data Archiving in the Digital Age: Strategies for Data Preservation and Sharing" (2013). *Anthropology Faculty Publications*. 1.

https://uknowledge.uky.edu/anthro_facpub/1

This Article is brought to you for free and open access by the Anthropology at UKnowledge. It has been accepted for inclusion in Anthropology Faculty Publications by an authorized administrator of UKnowledge. For more information, please contact UKnowledge@lsv.uky.edu.

Qualitative Data Archiving in the Digital Age: Strategies for Data Preservation and Sharing

Notes/Citation Information

Published in *The Qualitative Report*, v. 18, no. 24, p. 1-11.

The copyright holder has granted permission for posting the article here.



Qualitative Data Archiving in the Digital Age: Strategies for Data Preservation and Sharing

Lisa Cliggett

University of Kentucky, Lexington, Kentucky USA

Given the combination of recent mandates from funding agencies for data management plans and data sharing, and the explosion of data in digital form over the past two decades, it is time for the qualitative social science community to embrace digital archiving as an inherent component of research methodology. Archiving digital data ensures, at the least, that an individual scholar's data is preserved and accessible to the user many decades into his or her career. Digital archiving also has the potential to preserve for the broader scholarly community, the full range of social science knowledge far beyond an individual researcher's lifespan, or field site. However, the qualitative social science community has shown resistance to the archiving and data sharing movement. In this article I discuss the key debates around data archiving and sharing for qualitative research community, with particular attention to ethnographic data, and outline basic steps qualitative researchers can take as they begin to implement plans for digital archiving in their own research methodology. Keywords: Methods, Ethnography, Longitudinal Research, Anthropology, Qualitative Data Analysis, National Science Foundation, National Institutes of Health

Like it or not, the United States social science community is now obliged to include a “data management plan” in submissions for funding, thanks to the recent National Science Foundation (NSF) and National Institutes of Health (NIH) requirements addressing data preservation and sharing (in fact, Europe is far ahead on this path (c.f., Corti & Backhouse, 2005; OKFN.org; UKDA, n.d.). In this paper I address the longstanding concerns the qualitative social science community has grappled with about *how* data can and should be shared, and argue that it is now not only our duty under the new mandates, but also our duty as conscientious scientists, to take on the challenge of preserving and sharing *primary data*.

The concerns are both logistical and ethical. How does one go about managing vast sheets of crumbling records or hard drives full of digital files? Scholars could “scoop” the primary researcher’s findings. Restudies of primary data could disprove results and embarrass researchers. Study participants could lose their privacy, and the coherent ethnographic context inherent in primary researchers’ knowledge could be muddled. But with the new NSF and NIH policies emphasizing preservation and sharing of the raw data (not synthesized findings), it’s time for qualitative researchers to resolve those issues and take on the challenge of preserving and sharing *primary data*.

While data preservation and sharing is an issue for all social scientists, a specifically anthropological phenomenon is that a large number of scholars have, literally, a lifetime’s worth of data from one region. Wenner-Gren acknowledges this fact with its Historical Archives Program (HAP) emphasizing archiving and preserving research materials at retirement. With the growth of qualitative research in the 1960s, many disciplines, and especially anthropology, have reached a life cycle stage when scholars may consider archiving their data in their institution’s libraries or other respected archives for vaguely defined “access” by future generations (c.f., Kemper & Royce, 2002; Sanjek, 1990). In a few instances, scholars have passed their data on to a student or other collaborator. This

individualized type of data transmission, from one scholar to another, carries high risk for data loss (just as individualized systems of data backup are inherently risky).

In either case – whether library archive, or generational transmission - a key difficulty in making those data sets genuinely useful is capturing sufficient contextual depth so that a “newcomer” can gain meaningful understanding to work with the data, yet also protect study communities. Robust digital archives depend on detailed “metadata” that document general content of data sets and files (for use in indexing and finding guides), and also document inter-relationships between various data types and data content in ways that facilitate secondary users’ ability to work with and analyze data in a truly ethnographic style. While qualitative digital archives certainly exist (i.e., Online Archive of California: <http://www.oac.cdlib.org/>; Britain’s Qualidata: <http://www.esds.ac.uk/qualidata/about/introduction.asp>; Electronic Human Relation Area Files (eHRAF): <http://www.yale.edu/hraf/>), the ethnographic coherence of inter-related data types and content within a data set – exactly “the stuff” that most excites our ethnographic hearts and minds, is frequently the missing link.

In this article I outline the background issues and debates about qualitative data archiving and sharing, pointing to some of the most challenging tasks that scholars face when creating ethnographic archives. Because the current generation of scholars, and all who follow deal primarily with digital data, my discussion centers on the new technologies in which we store our data. Despite the challenges, and given the current digital foundation of qualitative work and data collection in the 21st century, I argue that as individual scholars, and disciplines, we must embrace archiving of our digital research materials to avoid, at the least, data loss through technological obsolescence, and ideally, create a profound resource for qualitative research, and science more broadly. A few early adopters might demonstrate to others that digital archives can indeed reveal vibrant ethnographic relationships and context in all their depth, while remaining attuned to fundamental concerns of confidentiality and protection of our research communities.

Debates and Challenges for Archiving of Ethnographic Data

Scholars have identified a number of barriers qualitative researchers cite when considering archiving and data sharing. Among these are concern for privacy, notions about data ownership and challenges for interpretation. I consider these issues below, and discuss one instance when such practices are more common.

Human Subjects / Privacy Concerns

Protection of study participants, sensitive data and confidentiality are some of the most frequently mentioned barriers to qualitative researchers archiving and sharing data (Carlson & Anderson, 2007; Moore 2010; Pienta, Alter, & Lyle 2009). However, depositing data in a digital archive does not translate to free, unmonitored access by the public.

In the past decade digital data archives have developed sophisticated systems for managing sensitive data and protecting study participants. All digital repositories recognize the importance of protecting study subjects, and consequently employ a number of strategies to ensure confidentiality, from highly restricted data set access, to research applications including Institutional Review Board (IRB) approval (see <http://www.icpsr.umich.edu/icpsrweb/ICPSR/curation/confidentiality.jsp> as one example). Some archives have also developed software to anonymize qualitative data, yet preserve context for researchers through code book identifiers (see, for example, Qualanon software, <https://www.icpsr.umich.edu/icpsrweb/DSDR/tools/anonymize.jsp>). With such software, the

search-and-replace capability allows for one level of identity protection; yet a code book of names/pseudonyms ensures the analytical integrity of data is maintained for possible future research needs (ICPSR, 2009).

Given that all social science data require attention to the protection of human subjects, all digital repositories must build systems for appropriate data access. Ultimately, researchers depositing data will work with archives to establish appropriate restrictions and software tools for particular data types.

Resistance to Data Archiving and Sharing

With technical and software developments promising useful and appropriate systems to manage confidentiality and ethics, qualitative data sets should increasingly be found in digital archives. However, compared with other social science data, ethnographic data sets are scarce in these repositories, raising the other most common barriers regarding data archiving and sharing: a perception of proprietary rights over data (AnthroDataDPA, 2009; Corti & Backhouse, 2005; Pienta, Alter, & Lyle, 2009), and disciplinary perspectives that privilege publication as the most appropriate form of data presentation and sharing (Kelty et al., 2008; Rabinow, 2006).

Scholars have debated the question of “who owns data?” at least since the early 1980s when the US congress passed the Bayh-Dole act protecting researchers’ patent rights associated with publicly funded research (Petsko, 2005). While the Bayh-Dole act addressed patent ownership, it did not clarify proprietary claims on raw data, leaving universities and research institutions to set their own policies of data ownership. Funding agencies have also weighed in on the question of data ownership, in so far as their policies require that data be disseminated and shared. The National Science Foundation’s policy on data sharing specifically identifies *primary* data: “Investigators are expected to share with other researchers, at no more than incremental cost and within a reasonable time, the primary data, samples, physical collections and other supporting materials created or gathered in the course of work under NSF grants.” (NSF, 2011, p. D.4.b). Furthermore, NSF states that “Grantees are expected to encourage and facilitate such sharing” (ibid). In effect, this policy argues that data produced through public funding belong to “the public.”

Despite the existence of recommendations and policies, differences in “world views” (in a cultural anthropology sense) exist between disciplines, which have resulted in scholars with quantitative data assuming that data belong to the university (or the public), while scholars with qualitative data believe their data belong to the individual researchers (Alter 2011). This is particularly apparent within anthropological sub-disciplines, when considering the explosion of the more “quantitative” data in archaeological, bio-anthropological and linguistic archives (see, for example: The Digital Archaeological Record [tDAR]; The Human Origins Data Base; The Open Languages Archive Community). While a few archives (i.e., the National Anthropological Archives and the electronic Human Relations Area Files [eHRAF]) contain collections of ethnographic archives, the majority of these come from retired or deceased scholars, rather than currently active researchers seeking to share data for scholarly inquiry now.

Perhaps these different epistemologies emerge through the presumed objectivity of data – “counting” (i.e., quantitative data) is considered objective data, while narrative field notes and conversations are data requiring at least one level of filter/interpretation as the researcher captures them through her own words (Carlson & Anderson, 2007). Of course this distinction is highly critiqued (e.g., Harvey, 1989; Spiro, 1996), but is not the subject of this article. Regarding “authorial rights” of primary researchers, current recommendations from archives encourage secondary users to cite primary researchers as “the author” of any data

sets used in research. In this way, the threat of being “scooped” is somewhat diminished and primary researchers receive some recognition for investment in creating the data set.

The other factor influencing qualitative researchers’ claims of data ownership may be the recognition that making sense of raw qualitative data requires the depth of contextual knowledge a researcher gains from “being there.” Thus, a secondary user would be incapable of interpretation (Carlson & Anderson, 2007). However, one goal of thorough metadata is to provide depth of context.

Recent discussions in the scholarly community about new technologies have celebrated the potential of open access publishing, and democratizing access to knowledge in the form of web publishing, blogs, “collaboratories” and other interactive web-based media (Braun & Whatmore, 2010; Fish, 2011; Kelty et al., 2008; Rabinow, 2006; SavageMinds.org, 2011; Suber, n.d.). These approaches to making publications more easily accessible are certainly welcome, and do revolutionize knowledge access. But they also speak to the notion that publication is the appropriate form of data dissemination, particularly for qualitative data which requires substantial interpretation by the primary data collector (see above).

However, even with new media outlets, “publication” still remains inadequate for the vast quantity and range of data that could be presented; not all words, images, recordings, videos, geospatial files, etc. will be included in a “publication.” Given the increasingly broad range of data types researchers collect, and given the increase in sheer quantity of data collected that new technologies facilitate, publication of any kind cannot remain the only, or best, path of data dissemination.

Project Inheritance/Longitudinal Studies

One instance when we see qualitative researchers – anthropologists in particular – joining the conversation about data sharing and data archiving is in reference to “project inheritance” as senior scholars consider grooming students or a new generation of colleagues to assume responsibility for project continuation (Kemper & Royce, 2002). While some longitudinal projects with multiple generations working in the same communities for decades are founded on a commitment to collaboration and data sharing (see, for example, chapters on the Gwembe Tonga Research Project [GTRP], the Kalahari Research Group, and the Tzintzuntzan Project in Kemper & Royce, 2002), very few have arranged for institutionalized data management and sharing across generations of researchers. In the project with which I am most familiar, Elizabeth Colson and Thayer Scudder have deposited their GTRP data in hard copy with their university archives, but priority for curation has been preservation, not data sharing (although that is under consideration for future digital archiving projects). More frequently, when data is passed on to a new generation, it is done so from individual to individual.

The risk of data loss from this kind of personalized transmission is extremely high, just as the risk of data loss from personal archiving systems for one researcher is high. With the growth in numbers of qualitative researchers engaged in research since the 1960s, there are an increasing number of scholars who have worked in one site over the span of their careers, and like the initiators of the GTRP and Tzintzuntzan Projects, these scholars may anticipate new generations of researchers joining the collaboration.

Finally, despite skepticism among some scholars about releasing data broadly, there are noteworthy examples of qualitative researchers engaging in collaboration and data sharing, though often through the mediation of a collaborative project (see, for example: Gravlee, Kennedy, Godoy, & Leonard, 2009, on the Tsimane Amazonian Panel Study; Hirsch, Smith, Wardlow, Parikh, Phinney, & Nathanson, 2009, on the five country "Love, Marriage and HIV" study; Lee & Biesele, 2002, on the Kalahari Project) or a restricted data

access web site (e.g., Tsimane Amazonian Panel Study <http://www.tsimane.org/>; Tsimane Health and Life History Project <http://www.unm.edu/~tsimane/>; Whole Village Project <http://wholevillage.umn.edu/index.shtml>). These examples of collaboration and data sharing help to normalize sharing and collaborating practices within the qualitative research community.

Transformation in Qualitative Perspectives and Practice

Given new technological and policy environments that encourage data archiving and sharing, combined with the exponential increase in longitudinal data from qualitative research projects over the past four decades, this is an important moment for development of data management systems and analytical capabilities for qualitative researchers. Even for scholars with data sets from a one year research project (including PhD research), agencies increasingly require data management plans that address data preservation beyond the primary researcher's use. By exploring and embracing possibilities for genuinely useful qualitative data archiving sooner, rather than later, scholars have a better chance to lay the foundation for what agencies, our universities and archiving institutions will require of us. In the process, increasing the number of qualitative data archives can begin transforming disciplinary notions about proprietary data and limitations for data sharing by demonstrating appropriate possibilities of archiving, sharing and collaborating.

Techniques in Archiving

In this section I discuss possible steps and techniques that all researchers can implement as they think about building a digital qualitative archive. In fact, these procedures have not yet been tailored specifically for ethnographic data, although I am currently experimenting with building a "best practices" system of digital archiving for ethnographic data (Cliggett, 2012). What follows are preliminary recommendations for qualitative archiving.

Keeping Data for the long term (the decision to archive rather than just backup)

With paper and hard copy data, scholars typically think about archiving long after collection and analysis. Archiving may occur when a scholar retires and deposits hard copy data in a university library, or when a project is long finished and the researcher needs shelf space for new material. With the majority of contemporary researchers keeping qualitative data in digital form (in text files, digital voice recordings, video, digital images, etc.), it becomes all too easy to stow a digital folder on a computer hard drive and move on to the next project, perhaps with the thought "I may come back to some of that material in a few years..."

This can be a dangerous choice depending on an individual's memory and organizational talents, and the technological characteristics of the particular computer and backup system. One of the biggest risks for the current era of scholarship is potential loss of data due to failures of technology. Hard drives crashing and poorly planned backup systems are easy to imagine. Even when backing up to separate "storage" spaces, such as external hard drives or the increasingly popular "cloud based" file storage systems such as "Dropbox," files remain vulnerable. Indeed, during a discussion in the "eFieldnotes" session of the March 2013 Society for Applied Anthropology meetings, audience members described how, earlier that month, Dropbox had a system glitch and many users discovered that their informal backup system had mysteriously deleted all of their files.

Far more reliable for protection of digital files are the new “digital repositories” emerging on most university campuses. Due, in part, to the new mandates from funders about data archiving and sharing, universities have begun to establish digital repositories where faculty can store their digital products, especially data, for both the short and long term. These repositories have as their primary purpose the preservation and protection of files, according to professional archiving standards. Combined with appropriate file formats (see below), these repositories (and digital archives in general) offer the best protection of digital files, both for short term backup needs, and for long term data preservation. Unless a storage option is identified as a preservation and archival system, digital files will remain at risk for technological catastrophe.

What’s less obvious in terms of digital vulnerability is the evolution of technology and proprietary software that, with time, makes old backup technologies and old digital files obsolete. As early as 1999, a colleague and I experimented with digitizing (scanning and using OCR- “optical character recognition” -) a portion of Elizabeth Colson’s field notes in order to explore possibilities for creating a fully digital qualitative database of the then 50 years of GTRP data. We saved files in an OCR format, storing them on “the standard” of the time – a 3.5 inch floppy disk. At that time, OCR software did a poor job of reading typewriter text (notes were typed using the old “field typewriter” common in the 1950-60s). We saved the files knowing that technology changes quickly and OCR would certainly improve over time. As often happens, a number of other projects came along before we returned to testing the OCR possibilities.

Now, 13 years later, we have a shoe box of 3.5 inch disks with files saved in 1990s proprietary software. Surely we could find technicians to free those files from their fossilized form, but it would require determination, time, and funding. In the current era of digital archives, and presented with the same challenge, the best option would be storing in *NON*-proprietary file types (i.e., open access format that will endure long beyond Microsoft’s and Apple’s dominance) in a digital repository. Most universities have joined the “repository movement” and have begun establishing a digital home for their research community’s data. Ensuring that data is accessible over time and technological revolutions, is the primary job of such repositories.

With a local digital repository on each campus, archiving digital data – whether for a few years until the next round of data collection, or for final closure on a ten (or forty) year project – should become as easy as making a copy of the data set to store in a file drawer. However, the ease of depositing data depends on a few key steps along the path of data collection and analysis.

The Basics: Privacy, Confidentiality and Anonymizing

Qualitative researchers often have a common misconception about depositing their raw data in an institutional archive: that such archiving is synonymous with open and unmonitored access, thus breaching our commitment to protection of study informants. As the section above on confidentiality explains, depositors set the terms for access to their data in consultation with the repository. Data can be embargoed (made completely unavailable) for a defined number of years, made available through an application procedure – including new IRB approval for the secondary users, and indeed, made more “open access” if the depositor chooses to designate the dataset as such.

However, best practices in any of these arrangements would include at least a basic level of anonymizing of individual and key identifiers (specific place names, etc.). For most of us, securing IRB approval means a commitment to confidentiality, typically including delinking identifiers. In practice, I suspect that ethnographers and others who work closely

with study populations, have a range of strategies to ensure confidentiality – perhaps using nicknames instead of proper names or using only first names in our field notes and interviews. Because ethnographers know their study populations with significant depth, it is virtually impossible for the researcher to anonymize to the extent that we cannot identify our own study participants. This sometimes confounds the more quantitative scholarly community. However, it is worth recognizing that if the researcher can identify individuals, others who know the community well may also recognize people. Before depositing data in an archive, researchers should ask the question: what would happen if someone who knows this community as well as I do, saw this data? The answers to that question can help determine the level of dataset restriction.

The possibility of anonymizing in ways more similar to quantitative datasets has not been fully explored for ethnographic data. However, the QualAnon application mentioned in the section on privacy has been used successfully with a range of other types of qualitative datasets. What remains to explore is how well QualAnon will work with ethnographic datasets that need linkage between files and data types in order to maintain the ethnographic context. My project to explore strategies for building an integrated ethnographic archive (Cliggett, 2012) will include testing the range of options that QualAnon offers with a goal of answering the question: will systematized anonymization still allow analysis of relationships and links between files and data sets, and over time? This issue – the effects of anonymizing ethnographic data for coherent usability – is crucial in promoting digital archiving for qualitative researchers. In order for such scholars to embrace digital archiving and sharing to its full potential, we need a system of protecting confidentiality AND preserving the knowledge context and value.

File Conversion to Durable Format

As mentioned above, the decision to deposit data in a repository ensures file preservation over time, despite technological change. However, most repositories require that depositors convert their proprietary software files (i.e., Microsoft or Apple file types) to durable non-proprietary digital formats so that data arrives in a stable form. This may simply require clicking the “save as” tab, and choosing “XML” for text files. For the majority of qualitative researchers, our most common data types are: text files, voice recordings, image files (both still and video); spreadsheet tables and possibly geospatial data. The standard durable formats for each of these file types is: Text files - XML format; Voice and video recordings - .WAV files; Image files- .tiff files; Spreadsheet tables- XML format; geospatial data – STDS. Most proprietary software offers a “save as” option that converts into these durable non-proprietary formats.

An important consideration in the file conversion process is what, if any, qualitative data analysis (QDA) software was used in working with the data set? Archiving of coded data (using one of the QDA packages, such as: Nvivo, MAXQDA, Atlas.ti) is another area for much needed experimentation. All three of the most popular QDA packages export to XML to some degree, and all show increasing awareness of the need for robust exporting options for durable formats.

However, we have very few examples of qualitative researchers testing the potential of maintaining the analytical power of coded data in the non-proprietary formats. Indeed, the scholarly potential of qualitative data archiving depends not simply on depositing datasets in archives, but preserving the knowledge already gained by coding / analysis in the XML format. With such preservation, secondary users can engage and build on that knowledge, by adding additional codes and engaging in greater analytical depth.

Metadata and the Archival Process

One benefit of using QDA software is that such digital coding actually produces one level of metadata that can be used in creating a digital archive. Metadata (the descriptive information *about* a file, a data set, or a project) is the foundation for usable archives. In the archival context, finding any dataset requires sufficient depth of metadata so that it appears when key search terms are included. Similarly, analytical power comes from sufficient metadata on the detailed file contents – (as are produced during coding with QDA software) – such as references to particular events (a funeral, healing ceremony, severe drought season, election, business transaction, etc.), particular families through time for case study analysis, particular time periods, and on. Importantly, in order to work with different file types from one project, metadata must serve as the link. Codes must be consistent across data types so that researchers can identify relationships between interviewees, images and spatial data, or regional relationships such as village residence and agricultural fields identified in maps, or topical themes addressed in different data types, such as interviews and field notes.

It is these kinds of linkages that capture contextual depth in qualitative data. While the original researcher may know these relationships without documentation, any subsequent user (whether a field researcher or an archive user) will rely on detailed documentation to make sense of any data set. The value of establishing extensive digital archives of ethnographic data, that facilitate data dissemination and sharing, will depend on the quality of the system of metadata accompanying the data set.

Current “best practices” of data archiving derive from the Dublin Core Metadata Initiative (DCMI) although some variation exists between archives/ repositories on the standard metadata fields, likely tied to the types of data sets in a given repository (DCMI, 2011; ICPSR, 2009; Van den Eynden, Corti, Woollard, Bishop, & Horton, 2011). Because of the qualitative depth of ethnographic data, archiving for such data sets will require modifications of the full DCMI system (see <http://dublincore.org> for their complete inventory of metadata fields).

Systems of metadata apply to “file levels,” such as one file of interview transcription, as well as higher order “data set” levels, such as a group of interviews focused on a particular interview guide, and the entire data set for a specific project. A challenging aspect of this documentation is the linking, through metadata fields, of the various types of data, so that, for example, it would be possible to extract all interviews with a given household, all photos that may relate to that household and any maps that might include geospatial data tied to the household.

While the DCMI list of metadata fields may appear overly systematized for qualitative researchers who work with their data in a more fluid process, recognizing that much of the metadata needed for archiving is already in our “head notes” (Sanjek, 1990) about our data, may make the process more palatable – it becomes a matter of writing down categories and tags, or including it during the data analysis process.

Conclusion

The realization that digital qualitative data is not indefinitely safe living on an individual computer or server is the starting point for qualitative data archiving. We must recognize that, as individuals and as disciplines, we are at risk of significant data loss if we rely only on data stored in a proprietary software format on our personal computers, university servers and cloud based file sharing systems. With that recognition, the next steps of anonymizing, coding (and related metadata creation) and saving as durable nonproprietary software formats become pro-forma.

With recognizing this need, the biggest challenge becomes internalizing archiving activities as part of our research methodology. While the current generation of faculty and senior researchers may have a majority of digital data, our students who are truly “digital natives” are the most likely population to fully implement digital archiving. But we must experiment with the range of options now, so that when those students plan their research agendas, they will have robust strategies on which to build.

Given the combination of recent mandates from funding agencies, and the explosion of data in digital form, it is time for the qualitative research community to begin digital archiving in earnest. At the very least, a scholar should ensure the preservation and usability of data over time, so that two decades after collecting the data, she could return to that dataset in preparation for a return visit to the same field site for a follow up project. On the other end of the continuum, archiving of digital data will lay the foundation for the broader research community to build on the knowledge one scholar has amassed. In either case, gaining an understanding of the basic steps and fundamentals of archiving for our ever growing disciplinary data sets should become a standard conversation, and practice, with our colleagues, and most importantly, with our students.

References

- Alter, G. (2011, August 2011). [Personal Communication. Director, ICPSR].
- AnthroDataDPA. (2009). Report from an NSF / Wenner-Gren supported workshop. May 18-20, 2009. Retrieved from <http://anthrodatadpa.org/addpa/>
- Braun, B., & Whatmore, S. (2010). *Political matter: Technoscience, democracy, and public life*. Minneapolis, MN: University of Minnesota Press.
- Carlson, S., & Anderson, B. (2007). What are data? The many kinds of data and their implications for data re-use. *Journal of Computer-Mediated Communication*, 12(2), 635-651.
- Cliggett, L. (2012). Strategies of data archiving for cultural anthropology, using Gwembe Tonga Research Project (GTRP) data. National Science Foundation - BCS- 1157418.
- Corti, L., & Backhouse, G. (2005). Acquiring qualitative data for secondary analysis. *Forum: Qualitative Social Research*, 6(2), Art 36.
- DCMI. (2011). Dublin Core Metadata Initiative. Retrieved from <http://dublincore.org>
- Fish, A. (2011, August 2011). Participation, collaboration and mergers. Retrieved from <http://savageminds.org/2011/04/12/participation-collaboration-and-mergers/>
- Gravlee, C. C., Kennedy, D. P., Godoy, R., & Leonard, W. R. (2009). Methods for collecting panel data: What can cultural anthropology learn from other disciplines? *Journal of Anthropological Research*, 65(3), 453-483.
- Harvey, D. (1989). *The condition of postmodernity: An enquiry into the origins of cultural change*. Oxford: Blackwell.
- Hirsch, J. S., Smith, D. J., Wardlow, H., Parikh, S., Phinney, H., & Nathanson, C. A. (2009). *The secret: Love, marriage, and HIV*. Nashville, TN: Vanderbilt University Press.
- Inter-university Consortium for Political and Social Research (ICPSR). (2009). Guide to social science data preparation and archiving: Best practice throughout the data life cycle (4th ed.). Retrieved from <http://www.icpsr.umich.edu/files/ICPSR/access/dataprep.pdf>
- Kelty, C. M., Fischer, M. M. J., Golub, A., Jackson, J. B., Christen, K., Brown, M. F., & Boellstorff, T. (2008). Anthropology of/in circulation: The future of open access and scholarly societies. *Cultural Anthropology*, 23(3), 559-588. doi: 10.1525/can.2008.23.3.559

- Kemper, R. V., & Royce, A. P. (2002). *Chronicling cultures: Long-term field research in anthropology*. Walnut Creek, CA: AltaMira Press.
- Lee, R., & Biesele, M. (2002). Local cultures and global systems: The Ju/'hoansi-!Kung and their ethnographers fifty years on. In R. V. Kemper & A. P. Royce (Eds.), *Chronicling cultures: Long-term field research in anthropology* (pp. 160-190). Walnut Creek, CA: Alta Mira Press.
- Moore, F. P. L. (2010). Tales from the archive: methodological and ethical issues in historical geography research. *Area*, 42(3), 262-270. doi: 10.1111/j.1475-4762.2009.00923.x
- National Science Foundation. (2011). Award and administration guide, Chapter VI- Other post award requirements and considerations. Retrieved from http://www.nsf.gov/pubs/policydocs/pappguide/nsf11001/aag_index.jsp
- OKFN.org (Open Knowledge Foundation). (n.d.). Retrieved from <http://okfn.org/about/>
- Petsko, G. A. (2005). Who owns the data? *Genome Biology*, 6(4). doi: 10.1186/gb-2005-6-4-107
- Pienta, A., Alter, G., & Lyle, J. (2009). The enduring value of social science research: The use and reuse of primary research data. ICPSR working papers *ICPSR Working Papers*. Retrieved from http://deepblue.lib.umich.edu/bitstream/2027.42/78307/1/pienta_alter_lyle_100331.pdf
- Rabinow, P. (2006). Steps toward an anthropological laboratory. *ARC Concept Note, 1*. Retrieved from http://anthropos-lab.net/documents/concept_labor/steps-towards-an-anthropological-laboratory/
- Sanjek, R. (1990). On ethnographic validity. In R. Sanjek (Ed.), *Fieldnotes: The makings of anthropology* (pp. 385-418). Ithaca, NY: Cornell University Press.
- SavageMinds.org. (2011). SavageMinds: Notes and queries in anthropology, *Savage Minds Anthropology blog*. Retrieved from <http://savageminds.org/>
- Spiro, M. E. (1996). Postmodernist anthropology, subjectivity, and science: A Modernist critique. *Comparative Studies in Society and History*, 38(4), 759-780.
- Suber, P. (n.d.). Why open access? *Open Access Anthropology blog*. Retrieved from <http://blog.openaccessanthropology.org/why-open-access/>
- Van den Eynden, V., Corti, L., Woollard, M., Bishop, L., & Horton, L. (2011). Managing and sharing data: Best practices for researchers. Retrieved from <http://www.data-archive.ac.uk/media/2894/managingsharing.pdf>
- UKDA (UK Data Archive). (n.d.). Retrieved from <http://www.data-archive.ac.uk/about>

Author Note

Lisa Cliggett is Associate Professor of Anthropology at the University of Kentucky. She has carried out ethnographic research in Zambia, southern Africa since the early 1990s, and is a member of the longitudinal Gwembe Tonga Research Project (started in 1956 by Elizabeth Colson and Thayer Scudder). With research specialties in ecological, economic and cultural anthropology, her most recent field work has focused on migration and environmental change in a Zambian national park buffer zone. Her publications include *Grains from Grass: Gender, Aging and Famine in Rural Africa* (2005, Cornell University Press), *Economies and Cultures: Foundations of Economic Anthropology* (with Richard Wilk, 2007, Westview Press) as well as articles in *American Anthropologist*, *Human Organization* and *Human Ecology*, among other journals and edited volumes. She is also the editor of the Society for Economic Anthropology's (SEA) new journal *Economic Anthropology*. She may be contacted at Lisa Cliggett, Associate Professor, Department of Anthropology, University of Kentucky,

211 Lafferty Hall, Lexington KY 40506-0024; Email: Lisa.Cliggett@uky.edu; Phone: 859-257-2796

Copyright 2013: Lisa Cliggett and Nova Southeastern University.

Article Citation

Cliggett, L. (2013). Qualitative data archiving in the digital age: Strategies for data preservation and sharing. *The Qualitative Report*, 18(How To Art. 1), 1-11. Retrieved from <http://www.nova.edu/ssss/QR/QR18/cliggett1.pdf>
