

# **ORGANIZE THIS!: Data management for anthropology in the digital age, preserving our evidence for future discovery**

2016 American Anthropological Association, Minneapolis, MN. November 2016

## CHAIRS/ORGANIZERS

Dr. Jerome Crowder, Institute for the Medical Humanities, U. Texas Medical Branch  
Dr. Richard Freeman, Smathers Libraries, University of Florida

## DISCUSSANTS

Lisa Cliggett (U. Kentucky)  
Deborah Winslow (NSF)

## PANEL ABSTRACT

As anthropology has moved into the 21<sup>st</sup> century, the transition from analog to digital has had profound effects on how anthropologists conduct their work and subsequently how their data are managed (recorded, analyzed, preserved). In this panel we consider the future of our research and scholarship, our evidence and discovery, and how we continue to engage with our data and productions. From requests for data management plans to archiving textual, visual and audio data, we consider how to deal with our anthropological digital evidence and the many accidents that can befall our research and the data we collect, analyze and store. This panel seeks contributors who can speak to their specific work regarding organization, preservation, metadata cores, access and retrieval (public and/or personal or selected group), archiving and policies at individual, institutional and federal levels.

Because anthropology focuses on the human engagement with others it is more important than ever to intelligently manage our electronic data. More than with analog materials, digital files can be widely disseminated and harder to control. Things that should not be made public too often are. This is simultaneously one of its many benefits, it can be easily shared with a few colleagues, one's subject group, or thousands of interested researchers and inquisitive members of the public. In fact, sharing data and findings with the public and data management schema is now a component in all government research grant applications. Understanding how to manage our digital data is not an afterthought, but central to the morals, and ethics of any practicing anthropologist and her success.

### **Digital Files and the future of anthropological data: ethics and organization**

Dr. Richard Freeman, Smathers Libraries, University of Florida  
Dr. Jerome Crowder, Institute for the Medical Humanities, U. Texas Medical Branch

As data management plans (DMP) are becoming increasingly required by more granting agencies (federal, State, Private), it is imperative that anthropologists not only understand the technical side of how to organize, share and store their data, but also the ethical implications of these actions, such as who will have access to this data and what will (what *can*) they do with it? Our paper reviews the work of several professional committees that have considered data management for anthropologists at the broadest levels, outlining technical schemas and considerations of what is possible and suitable for different kinds of anthropological research, touching on these technical and ethical issues. We will explore

different options available for researchers, including web-based and institutional services offered at many higher learning institutions. These services address different demands made on data by the granting agencies, from a plan for preservation to making the data more available to scholars and the public. No longer is “publically available” the simple publication of the work in a professional journal or monograph, instead we must embrace the options digital files provide researchers to disseminate their work so users can glean more from them. Open access, institutional repositories, and pre-publication versions, amongst other forms of data dissemination are now commonly considered as a part of many DMPs. We will then touch upon the technical side, how we organize our data, metadata schemes, storage, file types, etc. are all critical decisions when considering making the data usable into the near to distant future.

### **Digital Workflow in the Humanities and Social Sciences**

Dr. Smiljana Antonijević Ubois, Research Anthropologist, Penn State University

This paper presents findings of a three-year ethnographic study conducted at twenty-three academic institutions in the USA and Europe and funded by the Mellon Foundation. The study explored workflow of 196 faculty across the sciences, humanities, and social sciences, including anthropology, and it focused on the integration of digital technologies at all stages of a research lifecycle—from collecting and analyzing data, over managing and storing research materials, to communicating and sharing findings.

The results showed that digital technologies have different roles and levels of integration at various phases of scholarly workflow, and across academic disciplines. For instance, scholars in the humanities and social sciences (HSS) reported the highest percentage of lost and inaccessible files. Inadequate infrastructural, financial, and knowledge resources were commonly charged, but the academic reward system and scholars’ views on academic legacy also played a significant role, focusing their attention on the preservation of published materials versus data and work-in-progress content.

The results also indicated that privacy concerns in the HSS and complexity of multimodal fieldwork data influenced data management and storing practices, which were not adequately supported through the institutional technological and policy systems. Finally, the results showed that the majority of scholars in the sciences actively shared their research data, while the HSS scholars indicated opposite practice. The study found that in the HSS the character of objects of inquiry influenced data sharing practices, while academic status was the dominant factor across disciplines, with tenure-track faculty being more protective of their data than tenured scholars.

### **Hyperlinking anthropological analyses to its foundational fieldnotes: An example**

Dr. Janneke Verheijen, African Studies Centre, Leiden, the Netherlands

Cultural anthropologists often operate alone in the field, and, while always striving for “objectivity”, tend to generate data through essentially subjective processes. Data analysis too is commonly a solitary and subjective endeavor. Opening up access to raw data sets and offering insight into the process of data analysis adds to the trustworthiness of anthropological research findings. However, it also triggers new concerns.

In the digital version of my book *Balancing men, morals and money: Women’s agency between HIV and security in a Malawi village*, which is based on one year of in-depth

anthropological fieldwork, I have inserted hyperlinks that direct readers from my statements to the fragments of our fieldnotes on which these build (see <https://openaccess.leidenuniv.nl/handle/1887/21741>). This allows readers to weigh the validity of the conclusions that I have drawn and scrutinize my analyses. Furthermore, readers interested in a particular topic addressed briefly in the book can through these hyperlinks easily access the more detailed, richly contextualized diary entries or interview transcripts.

This is a novel way of opening up access to raw ethnographic data. In the proposed presentation I will share my reasons for and the practicalities of this approach, related to, among others, assuring sustainable future access in an era of rapid technological change. I will furthermore discuss the avenues that this new approach opens, as well as the (ethical) concerns that it raises – especially but not solely related to protecting research participants' privacy, and possible ways to deal with these.

### **Is Ethnographic Evidence Public or Private? Opportunities and Challenges to Contemporary Data Sharing Requirements with American Indian Tribes**

Dr. Sean Bruna, Department of Anthropology, Western Washington University

In 2003, the National Institutes of Health published the final requirements and guide for data sharing, explaining that data sharing is “essential for expedited translation of research results into knowledge, products and procedures to improve human health.” Many medical anthropologists and university IRBs, however, balk at the NIH requirement and either limit or completely embargo ethnographic data because of concerns regarding privacy and anonymity of research populations. As is argued, such data, and the evidences produced, are protected under federal HIPAA requirements, or at a minimum, anthropologically oriented ethical norms regarding anonymity. At the same time, many tribal nations, including the nation I partner with for health-based community-based participatory research, require that research partners submit all data collected to tribal archives at the conclusion of studies, often without specific guidelines regarding format or privacy restrictions. I argue that data sharing, new as it is to cultural anthropology, presents an opportunity to decolonize the discipline's ongoing trend of knowledge extraction by challenging our notions of control, ownership and management of ethnographic data. However, long-term sharing of ethnographic data in tribal archives presents unique ethical and infrastructural challenges to preservation and access, particularly when striving to accommodate tribal requirements in the context of IRB and HIPAA regulations. In light of these concerns, this paper discusses the vague data sharing requirements I agreed upon with a partnering tribe, disentangles specific ethical and infrastructural challenges to meeting the requirement, and presents policy solutions for future ethnographic data sharing with tribal communities.

## **The Past is Prologue: Preserving and Disseminating Archaeological Data Electronically**

Edward Schortman (Kenyon College), Jenna Nolt (Kenyon College), Ellen Bell (California State University, Stanislaus), Patricia Urban (Kenyon College)

Archaeological data are relevant to many questions posed in diverse fields. As sites are destroyed at rapidly increasing rates, benefitting from those findings often depends on reviewing the only traces of past human activities that survive: original field records produced during archaeological investigations. The sheer volume of these materials makes publishing them financially prohibitive, visiting archaeological repositories is costly in time and money, and handling paper records is potentially destructive. Making information about ancient socio-cultural processes available for use in diverse pursuits therefore requires digitizing and posting online all field records pertaining to archaeological investigations. The resulting archives greatly enhance: multidisciplinary research, including investigations in anthropology, history, and art history; innovative pedagogies in classes focused on the human past; and the ability of different stakeholders to engage their own histories directly through the archives. Transforming voluminous quantities of notes, drawings, photographs, and other paper records into searchable online collections requires collaboration among archaeologists, information scientists, and varied stakeholders to design protocols and test prototypes. There have been few initiatives launched to achieve these goals. This paper discusses ongoing efforts to digitize and post the results of archaeological research conducted from 1983-2013 across four valleys in northwest Honduras. The relevant files contain surveys of 941 sites, records pertaining to excavations conducted at 180 settlements, and analyses of over 1 million artifacts that, together, detail nearly 3,000 years of human accomplishments. Lessons learned in the course of these endeavors are highlighted along with the pedagogical, research, and outreach aspects of the project.