DESIGNING THE DIGITAL ARCHAEOLOGICAL RECORD: COLLECTING, PRESERVING, AND SHARING ARCHAEOLOGICAL INFORMATION

By Russell Alleen-Willems

A Thesis

Submitted in Partial Fulfillment

of the Requirements for the Degree of

Master of Arts

in Anthropology

Northern Arizona University

May 2012

Approved:

Kerry Thompson, Ph.D., Chair

Francis Smiley, Ph.D.

James Wilce, Ph.D.

License Information

Abstract

DESIGNING THE DIGITAL ARCHAEOLOGICAL RECORD:

COLLECTING, PRESERVING, AND SHARING ARCHAEOLOGICAL

INFORMATION

RUSSELL ALLEEN-WILLEMS

Archaeological digital data, like archaeological artifacts, are non-renewable resources that, once lost, are gone forever. Because digital data are so new in comparison to paper records, archaeologists lose data frighteningly often. First, this thesis summarizes my experience interning with Digital Antiquity, an organization specializing in preserving digital data. Second, this thesis details considerations in preparing, storing, and disseminating digital archaeological information. Finally, this thesis describes potential cultural, professional and educational concerns for users of digital archaeological repositories. As archaeologists create greater amounts of digital data, the digital curation crisis will grow. While a perfect solution has not yet been implemented, pioneering archaeologists have identified steps every archaeologist can follow to ensure that the fruits of their intellectual labors are not lost, while at the same time taking advantage of the unique properties of digital data to improve data and information sharing and use in archaeology.

Digital data are useful in ways that data on paper are not and cannot be. Digital data allow archaeologists to collaborate on large projects, communicate more effectively, and even reconstruct entire excavations. However, digital data are also far less stable than paper records. While a paper record may last well over a thousand years if kept in the right conditions, digital data are often unreadable in less than ten years, even when traditional preservation methods are used.

Digital storage is becoming the norm for archaeological publication, even though most archaeologists probably do not consider the downsides of digital publication. Moreover, many specialized archaeological data only exist in digital formats (e.g., laser scans, digital photographs, extensive databases) and archaeologists will (and have) lost these datasets because of faults in digital preservation. Digital data often represent the only record left after archaeologists excavate a site, and the loss of such valuable data is akin to bulldozing a site.

This thesis details my experiences during my internship at Digital Antiquity (an organization which specializes in data archiving), provides examples of other projects working on the digital curation crisis and gray literature problems, and finally discusses the needs that these organizations may not be considering fully in their plans. This information will provide a primer for archaeologists about what and who to consider when creating, publishing, and storing data and information in digital formats.

Acknowledgements

I would like to thank my parents for their loving support and encouragement in both my personal life and my schooling.

Thank you to the NAU faculty, especially Dr. Smiley and Dr. Wilce, my thesis committee members, and most especially Dr. Thompson, my thesis committee chair, whose edits and discussions about my thesis proved invaluable. Thank you to Dr. Miguel Vasquez for giving me a crash course in ethnographic research methods and helping me design my research project. Thank you to Adam Brin, Dr. Francis McManamon, Jim DeVos, and the rest of the staff at Digital Antiquity, as well as Dr. Mary Whelan of Arizona State University library for their guidance, advice, and friendly conversations. Thank you to Kim Ryan, Sierra Ryan, and John Langan for generously providing me with a home away from home for the summer, not to mention their friendly company! Thank you to Dr. Kelley Hays-Gilpin and the 2012 Post-internship class for their helpful feedback and comments.

Last, but far from least, thank you to my wife, Kelly Alleen-Willems, for going on the graduate school adventure with me.

Table of Contents

List of Tables

List of Figures

Chapter 1
Introduction: Digital Data in Archaeology

Archaeological digital data, like archaeological artifacts, are non-renewable resources that, once lost, are gone forever. Because digital data are new in comparison to paper records, archaeologists lose data frighteningly often. While paper records are more stable and easily preserved, digital records provide numerous advantages such as easier reproduction and dissemination that could significantly benefit the sharing of archaeological information. Pioneering archaeologists have developed steps to more effectively preserve digital data and have solved and improved many of the obstacles that digital archivists face. However, there are still many problems left to combat. Cultural and legal intellectual property concerns represent the primary problems left to tackle: what are the potential cultural, professional, and educational concerns about widely sharing digital archaeological data?

Attributes of Digital Data

Digital data are useful in ways that data on paper are not and cannot be. They allow archaeologists to collaborate on large projects, communicate more effectively, and even reconstruct entire excavations. In comparison, the only area where paper records are preferable is for preservation. While a paper record may last well over a thousand years if kept in the right conditions, digital data are often unreadable in less than ten years, even when traditional preservation methods are used.

Digital storage is becoming the norm for archaeological publication, even though most archaeologists probably do not consider the downsides of digital publication.

Moreover, many specialized archaeological data can only usefully exist in digital formats (e.g., laser scans, geographic information system data, computer aided drafting drawings, or extensive databases). If archaeologists are not careful, they will lose these datasets because of faults in digital preservation. While archaeologists still create paper records, specialized digital data often represent a more detailed and complex record of archaeological surveys and excavations. As archaeologists move towards collecting and recording more information in solely digital formats, the loss of such valuable data will become akin to bulldozing a site.

Thesis Outline

This paper details my internship experience at Digital Antiquity – an organization specializing in preserving digital data. This paper also provides examples of other projects working on the digital curation crisis and gray literature problems, details considerations in preparing, storing, and disseminating digital archaeological information. Finally, the paper describes potential cultural, professional and educational concerns for users of digital archaeological repositories. This information will provide a primer for archaeologists about what and who to consider when creating, publishing, and storing data and information in digital formats.

Chapter 2 provides an overview of the current state of data dissemination in archaeology and discusses the potential uses of a larger body of easily accessed archaeological information. Chapter 3 outlines reasons for preserving digital data. Digital formats offer archaeologists possibilities for recording, analyzing, and disseminating archaeological data that paper records simply cannot match. Electronic records provide

inexpensive and cost-effective reproduction and transport of huge reports and datasets. Archaeologists may also more easily combine electronic datasets to create aggregate datasets representing fieldwork across wide study areas. Some very useful electronic records, such as computer aided drafting (CAD) drawings, three-dimensional laser scans, and geographic information system (GIS) datasets lose utility and do not translate well into hardcopy formats.

With easier reproduction and transport, more archaeologists can access datasets and test past interpretations using new methods and theories. From a processual archaeology perspective, retesting and analyzing data allows archaeologists to more effectively support and refine interpretations and increases scientific validity. From a post-processual perspective, cheaper and easier access allows more people to interpret archaeological data, including populations who have been historically distanced from the research process, such as Native Americans. Archaeological theories may differ, but practitioners of different theoretical frameworks all value archaeological data as non-renewable resources that should be preserved.

Chapter 4 describes a number of the organizations and projects to improve data sharing and archival in the US and the UK. When undertaking digital data archiving projects of any scale, archaeologists may turn to these organizations for recommendations and advice. Chapter 5 focuses on the history, purpose, and staff structure of Digital Antiquity as well as details the projects I completed and my reflections on what I learned during the internship. Chapter 6 describes one of the major obstacles in computer-aided data discovery: that archaeological information rarely fits precisely into neatly defined "boxes," but that structured data is necessary for computer systems to "speak" to each

other. Chapter 7 covers legal and logistical obstacles in the US and traditional knowledge concerns about storing and disseminating archaeological information. Chapter 8 summarizes technological issues in data sharing and the Archaeology Data Service's "Guides to Good Practice" on a high level. Chapter 9 requests that archaeologists integrate plans for sharing and archiving "born-digital" data from their projects to ensure that their data is available to a wide number of publics, can be reused, and that recordation does meet the SAA ethic of preservation instead of quietly rotting away in a file cabinet or on a CD-ROM. As the title suggests, much work in digital data is left to be done and archaeologists need to experiment with new techniques, but need to prepare their data as best they can to be reused later and to stay aware of the different needs of archaeological publics.

Chapter 2
Locked Away and Forgotten: Gray Literature and Primary Data

In both the United States and the United Kingdom, federal laws mandate the conduct of archaeological work. In the United States, the National Historic Preservation Act of 1966 (NHPA), the National Environmental Policy Act of 1970 (NEPA), and the Department of Transportation Act of 1966 exemplify these laws (King 2004:9, Table 2.1). In addition to federal laws, many American states have more specific laws that mandate archaeological and cultural resource preservation. For example, the Revised Code of Washington requires a state permit for excavation of any archaeological site, even on private land (27.53 RCW, A.R.S. Sec. 41. Online). The preamble to the NHPA summarizes the primary idea behind these laws that "the spirit and direction of the Nation are founded upon and reflected in its historic heritage," (16 U.S.C. § 470(b)(1)) of which archaeological resources form a part. In keeping with this mission, Section 101(b)(3)(G) of the NHPA states that State Historic Preservation Offices (SHPO) must provide public information, education and training, and technical assistance in historic preservation.

Table 2.1 Summary of Cultural Resources Laws in the United States (adapted from King 2004:16).

Federal Law/Regulation	Purpose
Authorities that Deal with All Types of Cultural Resources	
The National Environmental Policy Act (NEPA)	Review and assess impacts of federal undertakings on the quality of the human environment (including cultural resources)
Executive Order 12898	Discourages adverse impacts that disproportionally affect environments of low-income or minority populations
Executive Order 12072	Requires cultural effects of siting agency projects in urban centers
Historic Preservation Authorities	
The National Historic Preservation Act (NHPA)	Review and assess impacts of federal undertakings on historic properties

Federal Law/Regulation	Purpose
The Historic Sites Act (HSA) of 1935	Established the National Park Service responsibilities for historic preservation; helped create the National Historic Landmarks (NHL) and Historic American Buildings Survey (HABS)
Executive Order 11593	Directs agencies to complete inventories of historic properties by 1974 and to treat eligible properties as if listed on the NRHP
Executive Order 13006	Requires agencies to give priority to using historic buildings in historic districts
Executive Order 13287	Directs federal agencies to review their compliance with section 110 and 111 of NHPA and report progress to the Advisory Council on Historic Preservation (ACHP) and NPS
Section 4(f) of the Department of Transportation Act	Prohibits any agency of the Department of Transportation from implementing a transportation project involving a historic property unless there is "no prudent and feasible alternative"
The Federal Tax Code	Encourages the preservation and rehabilitation of historic structures by offering tax credits to owners
Archaeological Authorities	
The Antiquities Act of 1906	Provides criminal penalties for disturbing archaeological resources on federal lands, also gives the Executive office the power to declare national monuments
The Archaeological Resources Protection Act (ARPA) of 1979	Provides stronger penalties for disturbing archaeological resources on federal lands, outlines a permitting process for archaeological excavation, and resolves ambiguities in the Antiquities Act
The Archaeological Data Preservation Act of 1974	Aka "Moss-Bennett" Act; directs all agencies to consider project impacts on archaeological, historical, and scientific data; authorizes up to 1 percent of a project's budget to identify and mitigate impacts
The Abandoned Shipwrecks Act (ASA)	Asserts U.S. ownership to shipwrecks in U.S. waters and then transfers control of wrecks to states
The Curation Regulations (36 CFR 79)	Establishes standards for curation facilities that care for federally owned artifacts, specimens, and related documents
Native American Cultural Resource Authorities	
The American Indian Religious Freedom Act (AIRFA)	Protects rights of Native American tribes to the free exercise of traditional religions; encourages agencies to consult when actions might affect the practice of traditional religions
Executive Order 13007	Directs agencies to avoid physically damaging "Indian sacred sites" on federal and Indian land and to allow tribal religious practitioners access to sacred sites
The Native American Graves Protection and Repatriation Act (NAGPRA)	Mandates that museums, universities, and other repositories with federal funding must inventory and repatriate Native American remains and associated burial artifacts to culturally affiliated groups

Federal Law/Regulation	Purpose
Treaties, reserved rights, the trust responsibility	Variable according to group, but rights not explicitly relinquished by a tribe in a treaty are retained and the U.S. is required to respect tribes as sovereign nations and attempt to make decisions that do not harm tribes
Historical Documents Authorities	
The Federal Records Act (FRA)	Directs agencies on how to manage records
Section 112 of NHPA	Directs agencies to preserve records of historic preservation activities in perpetuity

Archaeology in the United States

Professional archaeology in the United States is split between two main categories: academic archaeology and cultural resource management (CRM) archaeology. While exceptions exist, academic archaeologists are often able to choose the sites they study and take the time to carefully analyze and interpret the data they collect. Academic archaeologists may spend years focused on only one or a few sites in order to fully analyze and interpret the collected data. Academic archaeologists rarely have the resources to personally collect the broad datasets required to conduct far-reaching synthetic studies and so must limit interpretations of past human behavior to the specific areas and time periods for which they have the most detailed data. The goals of tenure, grant funds, and professional reputation encourage academic archaeologists to publish their interpretations, and sometimes their primary data, in peer-reviewed journals, edited volumes, professional meetings, and books.

Conversely, CRM archaeologists rarely select sites themselves and may only study sites that fall within the boundaries of a public land parcel or within the Area of Potential Effect (APE) of a project expected to adversely affect archaeological resources (Elston 1992: 40). In a best case scenario, a CRM archaeologist will document a site's surface expression and successfully convince project proponents to change their project

and avoid impacting the site. In these cases, the archaeologist will not excavate the site. Often sites are not avoided, however, and so CRM archaeologists must focus on mitigating archaeological damage by excavating sites slated to be destroyed by construction projects. CRM archaeologists often have only limited time to analyze and interpret the data gained from salvage projects, but will prepare a report and listing of the data for limited publication. Often, CRM publications only reach the federal agency overseeing the project and the project client. Agencies do archive CRM publications, but these archives are much harder to access and search than published books and peer-reviewed journals (Seymour 2010: 238). While some CRM companies and federal agencies do provide incentives for their archaeologists to intensely analyze, present, and publish in peer-reviewed journals on collected data, many companies and agencies do not. The lack of incentives to publish in peer-reviewed journals creates a backlog of CRM archaeological data and relatively few incentives exist to inspire corresponding efforts for more than minimal data analysis in CRM (Elston 1992: 40, Green and Doershuk 1998: 136).

Gray Literature

Archaeological work typically creates large bodies of survey and excavation data. In CRM, archaeologists communicate this data through written reports that detail the cultural resources found in the area that a project might adversely affect. These reports are typically under-published, with only a few physical copies ever produced. The scarcity of physical copies and the lack of centralized repositories mean that the results of such research are difficult to search and access.

Academic archaeologists reference CRM reports infrequently for a number of

reasons, including the difficulty in searching through CRM records for relevant information and because academic archaeologists attach a stigma to "gray literature" reports from CRM projects (Seymour 2010: 258). CRM "gray literature" represent an underutilized resource that all archaeologists may use to more effectively conduct new archaeological research by first making extensive use of the previous work done by other researchers. Schopfel defines gray literature as follows:

> Grey literature stands for manifold document types produced on all levels of government, academics, business and industry in print and electronic formats that are protected by intellectual property rights, of sufficient quality to be collected and preserved by library holdings or institutional repositories, but not controlled by commercial publishers i.e., where publishing is not the primary activity of the producing body [Schopfel 2011:17].

While the term "gray literature" may imply that the material is less rigorous or not up to professional academic standards, the term actually refers only to the method of publication and not to the quality of the information. Seymour (2010) argues that CRM reports display a range of quality from terrible to excellent that is similar to that found in academic literature. Academic archaeologists point to peer-review as a factor differentiating academic literature and gray literature. Seymour (2010: 260) notes, however, that many CRM reports also undergo extensive peer-review but stresses that peer-review is not always a guarantee of quality research. The data gray literature CRM reports contain often represent the only surviving information from archaeological sites that no longer physically exist. CRM reports record the bulk of the archaeological survey and excavations conducted in the United States and have potential to provide much more information about the archaeology of the United States than the relatively smaller number of academic reports (Green and Doershuk 1998: 121). In short, CRM archaeologists

possess a steady supply of opportunities and funding for archaeological fieldwork and data collection. Academic archaeologists possess resources and incentives to carefully analyze data and present interpretations in peer-reviewed publications. Wider publication and easier access to broad-scale CRM data sets would enable academic archaeologists to interpret past human behavior on much larger scales than currently possible and would encourage use of the data CRM archaeologists collect. Access to digital data and other CRM reports would allow CRM archaeologists to more easily compare data about sites in their project area to other sites in the surrounding region. In addition, increased access would allow archaeologists to more easily and efficiently conduct the background research necessary to prepare for archaeological surveys in an area. Digital data archival and dissemination offer the missing factors that would allow CRM and academic archaeologists to enter into a relationship that not only benefits archaeologists personally, but also improves the quality of archaeology as a discipline.

CRM Archaeology Research and Publication

CRM archaeologists often have to visit a number of separate archives when conducting background research in preparation for a project. These include government archives, such as those held by SHPOs or Bureau of Land Management (BLM) offices, as well as university and other archives. Often, the body of data any given archaeologist is able to collect only represents a small fraction of the total information about the area and may not include reports of all the previous fieldwork conducted in an area.

The problem with disorganized preservation and dissemination of data becomes even more critical, however, when archaeologists conduct excavation. Excavation

destroys the focus of archaeological study: the contextual relationship between objects, structures, and environments created by human beings. Once a specific unit or level is dug, the information is either carefully recorded or lost forever. In many cases, CRM archaeologists are the only people to record data about a site before the site is completely destroyed due to construction activities, leaving no opportunity for future excavations. As such, the excavation records, notes, images, datasets and sensory data are the only copy of a site left in existence. The destruction or loss of these records is a reprehensible act that deprives humankind of the information and other cultural values a site possesses. Indeed, destruction of records is, in essence, destruction of the site itself. According to the Archaeological Resources Protection Act of 1979 (ARPA), records and data on archaeological excavations must be preserved, making their loss through negligence illegal (16 U.S.C. 470cc(b)(3)).

Future Directions for Archaeological Data

As Eiteljorg (2004) notes, the economic climate of the early "twenty-teens" might limit the opportunities for field archaeology. Commercial and government enterprises are reducing the amount of construction in which they participate and universities are squeezed for funds. Correspondingly, if money for compliance archaeology and grant funded academic excavation is removed, "digging" through our large body of archaeological archives may be the best option left for continuing archaeological research. This economic reality could provide something of a boon to archaeology: the large body of underpublished gray literature, not to mention the incalculable body of unpublished primary data, likely contains a treasure trove of information waiting for archaeologists to

reanalyze using new research questions, theoretical orientations, and, eventually, new techniques.

Digitization of this largely untapped dataset would make the job of accessing and parsing through the material much easier than if the data remain in paper form. In addition, current archaeological data are being created primarily, if not solely, in digital formats (Archaeology Data Service and Digital Antiquity 2011). While more easily distributed and manipulated, digital data are at risk of loss through technological obsolescence unless archaeologists take special care to preserve them. Much like conserving a physical artifact in a museum collection, digital data need to be preserved using special techniques. Much of the archaeological data produced today are not being stored in sustainable formats, on sustainable media, or organized in such a fashion that even the same researchers may make sense of it after a few years.

If the proper steps are taken, digital data can be preserved far into the future. In addition, digital formats make it possible to quickly and cheaply share information with other researchers far more easily than data encoded in paper records. Instead of a researcher needing to travel to a specific physical archive in order to photocopy a few pages from an obscure report, the researcher could access that same report online through a digital repository such as the Archaeology Data Service archive or the Digital Archaeological Record repository. Digital archives and Web 2.0 archives, such as OpenContext, provide sophisticated tools that allow users to search for and discover relevant sources in a more efficient and complete manner than possible relying on visiting enough physical archives to achieve saturation (Kansa and Kansa 2011).

Tools like OpenContext enable efficient discovery tools that help ensure that

researchers receive the largest body of relevant data possible. Data researchers continually work to improve search functions through both refinements improvements to full text searches, which is similar to what one might expect on a web search engine like Google, and faceted search techniques, where data are ordered hierarchically in a series of taxonomical "drop-down" menus. Some researcher have even experimented with "tagging" systems, where users assign lists of keywords or "tag" that allow a single piece of data to appear in several different, sometimes conflicting, categories (Kansa and Kansa 2011:63). This ability for a single piece of data to show up in conflicting categories reflects debates in the archaeological community, where one site might be assigned to different time periods according to different archaeologists (Kansa and Kansa 2011:86).

The greater "discoverability" of data also has benefits for those who supply their publications and data to digital repositories. Kansa and Kansa note (2011:71), in fields besides archaeology, publications that are freely available online receive an overall increase in citation or those available solely in journals. Kansa and Kansa (2011:71) are working on methods to provide stable citations for online resources, not only as a way to provide citations for researchers and authors, but also as a way to track relationships and citations between publications, as well as datasets. In addition, a citation method for primary datasets and other non-publication resources would encourage further participation in data sharing. Kansa et al. (2011:209) note that citation of primary data can become a consideration of committees for academic tenure. CRM companies could benefit from higher citation counts of primary data that could serve as an indirect measure of the quality of a firm's data collection and reporting practices. When putting in a bid for an archaeological survey or mitigation project, a CRM firm could boast that

"data and reports compiled by our archaeologists are widely shared and cited, making sure that our clients' project comply with both the letter and spirit of the National Historic Preservation Act of 1966 and the Section 106 process."

The current state of archaeological information dissemination is inadequate considering the importance of preserving archaeological data. Archaeologists share information primarily through either formalized peer-reviewed publication or through informal sharing of publications and data sets through personal networks (Seymour 2010: 235). On the one hand, peer-reviewed publications ensure a level of quality for information but also significantly lengthen the time required to publish archaeological reports. Gray literature produced by CRM archaeologists, on the other hand, is largely only shared through personal networks of archaeologists, some of whom may differentially share information, or through tracking down the limited published copies that exist in management offices. The current state of information dissemination leads archaeologists to repeat archaeological surveys, lose information through obscurity, and generally conduct worse archaeological research work than might be done. Greater information documentation, discovery, and dissemination through digitization allows for archaeologists to avoid repeating fieldwork, consider a wider body of archaeological data for interpretation, and get their work noticed and cited.

The potential for sharing primary data, which is rarely shared by any means but through a personal network, allows for greater confirmation and testing of archaeological conclusions – a vital scientific practice largely unavailable to archaeologists until now. In the future, analysis of digital copies of primary data may form a large sector of professional archaeology. It also allows for diverse publics and interest groups to analyze

archaeological data. Opening up their data to public scrutiny may cause some archaeologists worry. Archaeological interpretations should always be held up for scrutiny, however, for processual archaeologists interpretations are always provisional and should be constantly re-evaluated and abandoned if disproved. Post-processual archaeologists argue interpretations should not be privileged and for them different interpretations of the same data present less difficulty (Hodder 1996:84). Archaeologists may be concerned that data may be used for purposes other than they which they intended, such as for looting or to support a tribe's claim to be federally recognized. Archaeologists should remove information that could lead looters or other groups to sites, such as New Age religions – who may inadvertently damage sites, from public versions of reports and datasets (ICOMOS Committee on Archaeological Heritage Management 2002). Otherwise, archaeologists need to accept that data can be put to multiple uses. Limiting data to a small, trusted population may "protect" the data from perceived misuse, but such restriction also precludes unanticipated novel and potentially useful purposes for the data as well. Similarly, archaeologists may have entered into agreements with employers or other groups to keep data private. Native American tribes represent sovereign nations who may have a prevailing interest in keeping some data private in order to preserve traditional intellectual property rights or sacred potency (see Chapter 7).

Archaeologists must respect the agreements they make, however, unless there is a prevailing reason to keep data private, like protecting Native American intellectual property. Archaeologists should refuse to enter into agreements that restrict their ability to disseminate both the collections and interpretations from their research. Some archaeologists may hold onto primary data if they believe that their plans for it may take

several years to come to fruition. The act of sharing a dataset does not preclude a researcher from continuing to work on that dataset. Sharing data does, however, allow other researchers to simultaneously conduct research on a dataset. This alternative is far preferable to cases where archaeologists keep data in their private collections, always planning to work on them "someday" but leave data unanalyzed because of other responsibilities.

Chapter 3
Who Cares? Why Preserve Digital Data

Archaeological databases, as symbolic representations of the physical archaeological record can be thought of, and interpreted, as texts. Archaeological databases, as texts, are subject to the same shortcomings, biases, and other divergences from an objective reality as are possible in any symbolic representation of the physical world. Greater sharing of, and access to, archaeological information benefits archaeologists by increasing the accuracy and rigor of archaeological interpretations. Access to archaeological sites and information also benefits non-archaeological communities by empowering communities with cultural information that has many uses. Organized intelligently, with full awareness of database limitations, digital archaeological databases provide a method for increasing access and dissemination of archaeological information to both archaeologists and other publics, while retaining control over legally and culturally sensitive information.

Differing Theories and Synthesis

Archaeological theory primarily consists of two prevailing schools of thought: processualist theory, particularly that of Binford (1965; 1968), which classes archaeology as a field of objective scientific inquiry into past human behavior, and numerous post-processualist theories and critiques (Shanks and Tilley 1987, Hodder 1996), which assert that archaeologists and their interpretations are inherently and inescapably biased. While the schools of thought disagree fundamentally about the purposes, methods, and roles of archaeology, and even about the nature of knowledge itself, both schools must utilize the

same archaeological record as a basis for interpretation. While I disagree with the more extreme post-processualist assertions (Shanks and Tilley 1987), I do agree with the post-processual critique that archaeologists are inherently exposed to the political and social currents of the settings in which they operate, however, the archaeological record and, to a lesser extent, archaeological data, exhibit a mitigating influence over these biases since interpretations must relate back to the material record. Binford (1980:4) notes that "the archaeological record is at best a static pattern of associations and covariations among things distributed in space. Giving meaning to these contemporary patterns is dependent upon an understanding of the processes which operated to bring such patterning into existence." Like Binford (1980), I assert that archaeologists must try and refine their understanding of archaeological and cultural practices in order to give meaning to the patterns found in the archaeological record. This process of continual refinement means that archaeologists must continually re-examine past archaeological data under the lenses of both new theoretical frameworks and the continually expanding body of archaeological data from more recent and current fieldwork. The archaeological record, however, does not change, only interpretations of the record change. Accordingly, the record and derived data ground archaeological interpretations, since the interpretations must always adequately explain the patterns found in the archaeological record. Hodder (1996:172) notes that the act of data collection, while done using particular theoretical frameworks, is largely independent of higher level theoretical analysis. Hodder also proposes that, so long as the original theoretical framework used during collection is known, the data collected can still be used in other frameworks. Currently, however, many archaeologists worldwide lack the access to the wide and varied body of

archaeological data required to make the best use of the mitigating influence or to re-examine data using new frameworks. I believe that more effectively organized archaeological databases, planned for from the very beginning stages of research and fieldwork, provide the key to unlocking a door separating archaeologists from utilizing the past body of archaeological data for both processual and post-processual re-interpretation.

Putting archaeological data into any sort of structure, including databases, consists of interpreting the archaeological record and translating it to a different form. Archaeologists cannot objectively record facts from the archaeological record and turn them directly into data (Hodder 1996:169). In both computer science and archaeological theory, translation of the primary record into a secondary representation risks the data becoming corrupted. Archaeologists must strive to be aware of how the organization and design of any databases affects, changes, and constrains the form of the data as well as subsequent interpretations (Oppel 2010; Bourdieu 1977; Presland 1984).

Hodder notes:

> subordinate groups who wish to be involved with archaeological interpretation need to be provided with the means and mechanisms for interacting with the archaeological past in different ways. This is not a matter of popularizing the past, but of transforming the relations of production of archaeological knowledge into more democratic structures [Hodder 1996:186].

These archaeologically disenfranchised publics and subordinate groups include Native Americans, the many publics interested or invested in archaeological research outcomes, and essentially anyone who is not a professional archaeologist. Native American groups benefit legally and politically from access to archaeological data. Archaeological evidence has been used to uphold treaty claims to traditional resources, such as fishing

rights in Washington State (Trosper 2002). Hodder (1996) also discusses providing subordinate groups with access and technical knowledge about archaeology so that the dialogue in interpreting the human past becomes multivocal and different groups interpret data into knowledge that, while restricted in variety by the mitigating influence of the material record, becomes valuable to groups in the present. Archaeologists and the archaeological record have also suffered from the general lack of knowledge among non-archaeologists about the theory, practice, results, and valuable uses of archaeological research (Green and Doershuk 1998: 141).

Protecting Sensitive Data

Digital databases allow users to view the same data in different ways (Oppel 2010). Databases such as Digital Antiquity's Digital Archaeological Record (tDAR), allow members of the public to view site reports with all locational information removed, following the legal requirement of ARPA that site locations be protected to prevent damage and looting. As Richards and Ryan (1995:6) explain, with proper precautions archaeological data can be secured and released only when and in the format desired by those responsible for protecting archaeological information and sites.

Theoretical Encapsulation

To summarize, the theoretical orientation of my internship thesis is as follows: An objective reality exists. This objective reality contains a theoretically knowable past (Binford 1965). The archaeological record, though modified and made incomplete by various cultural and natural formation processes, provides traces of human behavior

archaeologists can use to write representations of the knowable past (Binford 1968; Schiffer 1972). Archaeologists are theoretically capable of accounting for and controlling for many, if not all of the distortions caused by formation processes, even though the task may seem near-Sisyphean at times (Binford 1980). A wider symbolic representation of the archaeological record, reanalyzed periodically by archaeologists using various (new, old, and yet to be created) theoretical positions and by comparison with the continually expanding body of archaeological evidence is key to accounting and controlling for interpretive biases (Reilly and Rahtz 1992:9). Well-defined, yet evolving, standards of planning, organization, preservation, interoperability, and open access of archaeological digital databases represent a pathway towards creating, preserving, and disseminating a digital archaeological record.

Chapter 4
The Players: Organizations in the US and the UK

A number of organizations are actively working to preserve and disseminate archaeological data and information. Here, I describe a few that are based in the United States and the United Kingdom, along with how they vary in their purpose and methods.

Digital Antiquity

Digital Antiquity is a non-profit organization located in Tempe, Arizona. An archaeologist at Arizona State University (ASU), Dr. Keith Kintigh, founded Digital Antiquity. While Dr. Kintigh is still actively involved with Digital Antiquity, Dr. Francis McManamon, the former head archaeologist for the Department of the Interior and the National Park Service, serves as Executive Director. Dr. McManamon oversees a team of technologists and archaeologists, including Adam Brin, the director of technology; James DeVos, a computer programmer; and a team of doctoral students in archaeology, who work as data curators.

The team at Digital Antiquity created a digital repository known as the Digital Archaeological Record, or tDAR for short. Focused primarily on the preservation of information and data, tDAR archives submitted information in the original digital formats (e.g.,, Microsoft Excel spreadsheets, Word documents, Adobe PDFs) to preserve the original form of the data as closely as possible. tDAR also migrates information to more stable preservation formats (e.g.,, text documents, archival PDFs, comma delimited text files), which the database also archives. Preservation formats ensure users can easily access information in the future, even if the original submission format becomes

incompatible with future hardware and software. Digital Antiquity is constantly updating and upgrading their system to accept new kinds of data which developers deemed too complicated for the first version of tDAR and have plans to add geospatial data in the near future (e.g.,, Geographic Information System – GIS – Shapefiles).

Alexandria Archive Institute

The Alexandria Archive Institute, based in San Francisco, California, runs a different kind of online archaeological information resource. Called OpenContext, the system differs from a digital repository in several ways. First, OpenContext stores their files sustainably, but does so through a partnership with the California Digital Library, who oversee and maintain the storage of OpenContext's files. OpenContext is focused on primary data publication and dissemination, rather than acting more like a traditional library and repository, which distribute finished reports. Using a motto of "one webpage per potsherd," OpenContext stores individual records for artifacts, excavation units, and archaeological levels from archaeological investigations (Kansa and Kansa 2011). For each record, data curators produce and store metadata, descriptive, and contextual information (such as spatial location) using extensible markup language (XML) format. The XML format is based on the ArchaeoML language (Schloen 2001) and allows both humans and machines to interpret the data. Machines can then conduct searches that link together all information in OpenContext that matches certain XML criteria, such as a specific date range or artifact type. This special formatting of metadata is an integral part of building a "Semantic Web" of information through the internet (Kansa and Kansa 2011).

Past Organizations in the United States

Digital Antiquity and the Alexandria Archive Institute are not the first digital archaeological archives in the United States. Two other initiatives began with similar objectives of preserving and sharing archaeological reports and data, but were unable to overcome financial and societal obstacles. Harrison Eiteljorg II and the Center for the Study of Architecture (CSA) created the Archaeological Data Archive Project (ADAP) in 1996. Like Digital Antiquity and the Archaeology Data Service; ADAP was a digital repository organization designed to preserve archaeological reports and datasets in digital formats. ADAP also realized the need to not only store the original data format, but to also translate data into sustainable formats. Unfortunately, Eiteljorg and the CSA were forced to discontinue ADAP in 2002. As for so many projects, lack of funds was a critical factor in ADAP's closure, but the project also suffered from not assembling enough material from archaeologists to become an indispensable resource (Huggett 2006). Managers of other digital archaeological repositories will need to work diligently and generate investment by the archaeological community if they want to avoid ADAP's fate.

Archaeology Data Service

The Archaeology Data Service (ADS), based at the University of York in the United Kingdom, is the most successful of the archaeological data repositories. Founded in York during 1996, the ADS preserves and disseminates full archaeological reports and datasets, as well as metadata information that help archaeologists locate materials not curated by the ADS. The ADS “ArchSearch” online tool provides interested

archaeologists and members of the public all over the world access to full reports and datasets from sites and monuments throughout the United Kingdom. Many of the reports available through the ArchSearch database are "gray literature" archaeological reports from private firms that normally receive very limited print publication.

Between 1998 and 2002, the Archaeology Data Service compiled and published hardcopy guides for archaeologists on best practices in creating and archiving digital data. These original "Guides to Good Practice," included information on preserving data from aerial photography, excavation, geophysical survey, GIS, CAD, and even virtual reality as well as digital versions of textual reports, spreadsheets, databases, and images (Archaeology Data Service and Digital Antiquity 2011). As part of my internship, I helped Digital Antiquity collaborate with the Archaeology Data Service to update the Guides to Good Practice to include information specific to conducting archaeology in the United States. The Archaeology Data Service publishes updated versions of these guides through their website.

Conclusion

Allowing archaeologists to examine original datasets with different methods and theories can help researchers support valid interpretations and discard invalid interpretations. Archaeological theories give many reasons for preserving digital data from archaeological projects. Processualist archaeologists should be pleased with the opportunities for scientific rigor that long-term preservation, broad access, and quick dissemination of datasets allow. Similarly, post-processualist archaeologists can adopt digital preservation and dissemination techniques to construct multiple interpretations

from data. Because of the easy dissemination digital formats allow, interpretations can come from many different audiences, including indigenous communities and other traditionally silenced voices in archaeology.

Chapter 5
Digital Antiquity Internship Background and Report

My internship focused on the use of technology and the internet to preserve and share archaeological data: both primary data and analytical reports. In the summer of 2011 I began a nine week internship working for Digital Antiquity, an organization based out of Arizona State University. I worked with Adam Brin, Digital Antiquity's director of technology, on projects developing best practice guides for digital preservation as well as projects that used technology to preserve data.

Digital Antiquity Organization History

In 2004, the National Science Foundation (NSF) funded a workshop on the integration and preservation of structured digital data from archaeological investigations (McManamon and Kintigh 2010:38). The workshop participants reported that archaeologists needed to create an infrastructure for preserving, accessing, and integrating archaeological information in order for archaeology to advance long-term, scientific understandings of human history. The NSF later granted $750,000 to Keith Kintigh and a group of Arizona State University archaeologists and computer scientists to create a prototype digital repository for archaeological data. The Andrew W. Mellon Foundation supplied an implementation grant that established an independent organization, Digital Antiquity, to further develop and maintain the repository, then known as the Digital Archaeological Record (tDAR). In November 2009, Dr. Francis P. McManamon, formerly the Chief Archaeologist of the National Park Service and Departmental Consulting Archaeologist for the Department of the Interior, joined Digital

Antiquity as the full-time Executive Director.

Digital Antiquity Organization Structure

Currently, Arizona State University houses Digital Antiquity in the Hayden Library. In addition to the Executive Director, Dr. McManamon, Digital Antiquity's staff comprises a Director of Technology (Adam Brin), one full time software engineer (James deVos), several digital data curators (Shelby Manney, Scott Thompson, Hannah Rivera and Josh Watts) and an administrative support person (Kirsten Clary). Digital Antiquity's Board of Directors consists of eleven members drawn from academia and commercial organizations; Jeffrey Altschul from the SRI Foundation and Statistical Research, Incorporated; Fred Limp from the Center for Advanced Spatial Technology at the University of Arkansas, which ran the National Archaeological Database effort; and Julian Richards, from the Archaeology Data Service. Digital Antiquity also utilizes an additional eleven member group of Science Advisors, and includes Eric Kansa, the lead software developer for the OpenContext project at the Alexandria Archive Institute (Digital Antiquity 2012).

Digital Antiquity Projects

Digital Antiquity operates a number of different projects in digital report and data archival and publication. The primary project, the Digital Archaeological Record, consists of a robust computer database and archival system with a front-end interface users access via the web. Secondary projects include advertising and promoting use of tDAR through presentations at professional conferences, funding digitization projects, writing scholarly articles, and advocating for the preservation and dissemination of archaeological

information.

Digital Antiquity developed tDAR to meet two needs. First, tDAR provides a system for accessing and using written archaeological reports and other information such as images, tabular data sets, and even 3D laser scans. tDAR contains computerized search functions that allow users to locate archaeological information from previously underutilized and obscure sources such as cultural resource management projects. The references included in underpublished cultural resource management reports or "gray literature" can help users become aware of the existence of other underpublished information kept in physical or digital archives by regional U.S. Bureau of Land Management, National Forest Service or State Historic Preservation Offices (McManamon and Kintigh 2010). Second, tDAR provides long-term preservation of data from archaeological investigations (Digital Antiquity 2012). Digital Antiquity stores data in both the original submission format (e.g., .xls for Microsoft Excel spreadsheets, .doc for Microsoft Word documents, .jpg for images) and prepares a preservation copy of the data in a form that will remain readable in the future (e.g., comma separated values encoded in plain—human readable— text) (Digital Antiquity 2012).

Over the past two years, Digital Antiquity staff advertised tDAR to a wide number of professionals in federal service, state and tribal historic preservation offices, cultural resource management firms, and academia by offering presentations and papers at a number of professional meetings (McManamon 2011:4). These presentations helped to raise awareness among professional archaeologists about the existence of the tDAR repository and demonstrated benefits of repository use.

In 2011, Digital Antiquity began offering grants of between $2,000 and $10,000 for

projects to make unpublished and underpublished archaeological reports, images, data sets, and other information available through digitization. The fifteen funded grants cover both academic and cultural resource management projects and will make hundreds of archaeological reports and datasets more easily accessible (McManamon 2011:9).

Digital Antiquity has offered advice to the National Science Foundation, the National Endowment for the Humanities, and the White House Office of Science and Technology Policy about how publicly funded archaeological projects and offices should preserve and disseminate their digital data (McManamon 2011:4). In addition, Digital Antiquity works closely with the Archaeology Data Service at the University of York to compile known best practices to ensure archaeologists do not lose digital information due to changing computer hardware, storage formats, or poor documentation. The guides help archaeologists make wise decisions and carefully plan the collection, use, and preservation of digital archaeological data from projects.

Internship Experiences

During my internship with Digital Antiquity, I worked on projects designed to collect, share, and exemplify best practices for preserving data in digital formats. Working towards that goal, Digital Antiquity worked with the Archaeology Data Service (a UK organization) to revise an online document: The Guides to Good Practice, on digital archiving in archaeology. I spent days reading through the draft version of the Guides and offered edits and suggestions for changes. My preceptor, Adam Brin, and Digital Antiquity's executive director, Dr. Francis McManamon, told me that proper archival is crucial to preventing digital archaeological data from degrading and becoming unusable. Indeed, unlike paper copies of archaeological data, which could last potentially for hundreds to thousands of years, digital

copies of data routinely became unreadable in decades or years (Francis McManamon, personal communication 2010).

My original internship plan included gaining training and experience using the computer scripting language Javascript to create visualizations of archaeological data. Interactive visualizations are a unique benefit of digital data that allow users to quickly and easily understand large amounts of information. While I did receive training and experience in using Javascript through access to tutorials available through Arizona State University's online library, the visualization Adam Brin requested (an interactive chart of the data types available in the tDAR archive) was beyond my skills to produce as quickly as my preceptor wanted. I was, however, able to help Brin create another visualization: a world map showing the point of origin for different datasets, not by writing Javascript code, but by looking through and modifying the extensible markup language (XML) tags used to encode geographic data for use by computers.

Digital Repository Funding Research

During the first week of my internship my preceptor asked me to write a report comparing different pricing strategies for digital repositories in other disciplines. At the time, the Arthur Mellon Foundation had invited Digital Antiquity to apply to renew a grant they had received. As part of the next stage of funding, Digital Antiquity needed to calculate what price to charge data depositors in order to cover the cost of ingesting, storing, and migrating deposited data indefinitely.

I researched and compared costs for different digital repositories in disciplines like biology, chemistry, and physics. I found that, while many fields are developing digital tools for data sharing, few projects shared Digital Antiquity's focus on preserving digital data for as long as possible. Of these organizations, the Archaeology Data Service in the UK was the

most comparable in purpose and scope. The Archaeology Data Service factored into their cost predictions an interesting fact about computer hardware: that computer harddrives historically display a trend of exponentially increasing in capacity while exponentially decreasing in cost. Because of this trend, an aspect of "Moore's Law" which is named after Gordon Moore, the co-founder of Intel who first described the trend (Moore et al. 1998), the cost of storing data quickly become inconsequential. Human costs, however, do not decrease with time, but tend to increase. While the cost of storing data may all but disappear, the cost of human effort to migrate data to new, compatible formats increases. The Archaeology Data Service mitigates this issue by requiring that submitters encode data in "preservation formats" that curators can be easily migrate to usable formats whenever a person wishes to use it. By not having to constantly update the format of a piece of data, the cost of preserving data becomes inconsequential, and humans only have to expend noticeable effort only when someone wants to use a piece of data, limiting human effort to only those pieces of data that people access. In addition, if data translation processes become automated, the automation also minimizes the human cost for data migration.

Digital Repository Metadata Comparison

I also compared the "metadata" systems different digital repositories used. Metadata is descriptive information about a resource including an identifier for a file, who created the file, the date the creator made the file, the topic of the file, etc. The literature often refer to metadata as "data about data" and well-structured metadata make discovering relevant data easier, similar to a card catalog in a library (Richards et al. 2011: 33). As discussed in Chapter 6, different metadata systems can affect interpretations and uses of archaeological data as they introduce a source of bias. Well-designed metadata systems are flexible enough to allow users to create their own categories that precisely describe data but structured enough to

enable computer algorithms to identify similar resources.

Guides to Good Practice

One of the primary deliverables of my internship was to offer comments and help revise three sections of the Guides to Good Practice. Through my internship I read through the Guides in their entirety and offered comments on each of the chapters. Although most of my comments were copyedits, my preceptor felt my contributions merited a contributor credit when they print the Guides in hardcopy.

The Guides offer advice and best practices on storing digital data from myriad archaeological research products: the common documents, databases, spreadsheets, and images most projects deal with, but also aerial and marine survey data, laser scans, close-range photogrammetry, GIS, CAD, and even virtual reality data (Archaeology Data Service and Digital Antiquity 2011). Through reading the Guides, I learned much about the considerations that archaeologist must make to ensure digital data are not lost prematurely; information I share in more depth in Chapter 7.

Grantee Interaction

An unexpected bonus of my internship with Digital Antiquity was that I received a little experience working on the "giving" side of a grant. As part of their efforts, in 2011 Digital Antiquity selected several organizations to receive grants for projects digitizing archaeological information for deposit into tDAR (McManamon 2011:7). The program not only built up the tDAR collection, making it a more useful source of information, but also advertised the existence of the repository and familiarized several organizations with the process of preparing and uploading data into tDAR. I was responsible for helping Shelby

Manney, one of Digital Antiquity's data curators, process grant paperwork through Arizona State University's grant office so that grantees could receive their funds.

Kennewick Man Digitization

One of my primary internship goals was to gain experience digitizing and preparing data for long-term preservation. Francis McManamon, Digital Antiquity's executive director, identified a dataset that was frequently used by archaeologists, but was in danger of disappearing because of technological obsolescence. As the former National Park Service head archaeologist, Dr. McManamon helped prepare several reports on the famous Kennewick Man remains from Washington State. The NPS published these reports electronically on an NPS website in May 2004 (http://www.nps.gov/archeology/kennewick) in order that publics might access them. Recently, NPS began migrating all their webpages to a new content management system and Dr. McManamon became concerned that the Kennewick Man reports website might not translate well to the new system. In order to ensure that the reports would be available in a similar format, Dr. McManamon had me redigitize the reports as a series of PDF documents. After I converted the reports to PDF files, I uploaded the documents to the tDAR repository and created descriptive metadata so that users can easily find the reports by searching tDAR with related terms. In addition to the original NPS webpage, the Kennewick Man reports are now also available at http://core.tdar.org/project/6325 where Digital Antiquity will preserve the reports indefinitely.

Questions and Concerns I had during my Internship

During the course of my internship, several experiences I had made me question how

Digital Antiquity could best serve different audiences besides just other archaeologists. Having large amounts of archaeological information available online to anyone with a web browser has great potential to enhance not only archaeological research, but also public education about archaeology, and a disruption of archaeological authority that allows multiple interpretations of archaeological information (Fuka 2009) . When I was re-digitizing the Kennewick Man reports, I was disturbed to notice that one of the reports (Walker, Larsen, and Powell 2000) contained images of the remains themselves. While not unusual for this kind of report, I was still a little surprised that pictures had been so readily available online for several years. Knowing that many groups hotly contest the Kennewick Man disposition and some indigenous groups have taboos concerning images of human remains, I wondered if tDAR could provide a disclaimer warning people that a report or other data displayed human remains. Another software project, Mukurtu (http://www.mukurtu.org/), provides users tools to control sharing culturally sensitive information and helps users design the system to conform to local cultural protocols. This experience, as well as others, made me question who, besides archaeologists, could be interested in having access to digital archaeological data, and what needs they might have. In Chapter 6, I explore this topic and outline an ethnographic research plan to investigate the needs of different potential digital archaeological data user groups.

Chapter 6
Square Pegs and Round Holes: Conceptual Reference Models, Site Records, and the Semantic Web

Standardized terms provide a way to more easily discuss complex concepts but sometimes even simple concepts require detailed description to precisely discuss. Archaeologists have debated the merits of both standardized terms and detailed description in archaeology's past and this debate continues in the present. Ambiguity in terms causes real problems when archaeologists try to use and combine different archaeological datasets. However, archaeologists can minimize the differences between terms in order to use terms as shorthand to talk about complex concepts. Digital dissemination of archaeologists' primary data offers immense potential for conducting broad analyses using diverse datasets. However, terminological ambiguity presents immense difficulty for combining multiple datasets in meaningful ways. Eiteljorg is pessimistic that archaeologists can ever resolve enough terminological ambiguity to successfully combine datasets. Other authors, such as Limp (2011: 278-279) and Kansa and Kansa (2011) believe that the problem is difficult but not insurmountable.

Data sharing would allow archaeologists to collaborate on projects, but allow data sharing across projects. For example, an archaeologist could compare ceramic measurements the archaeologist collected from one site to ceramic measurements other archaeologists collected from several contemporaneous sites. Comparative datasets would allow the archaeologist to make statements about how ceramics at a single site compare with regional ceramic attributes. Archaeological terms rarely conform precisely to any particular standard, however, and different archaeologists may have recorded entirely different attributes, used different terms to mean the same attributes, or used the

same terms to mean completely different attributes. The archaeologist would be making a large error in uncritically combining datasets that use different terms. Although the differences in the use of a term may be slight, even small differences can be critical. Terminological ambiguity makes combining disparate datasets into large aggregate and meaningful datasets very difficult.

Eiteljorg (2008:214) notes that controlled vocabularies may work for some aspects of archaeology, such as lithics analysis, where terms are relatively few and precisely defined. However, for most archaeologists standardized vocabularies limit their ability to tailor research projects for specific environments and to answer specific research questions. Wendrich (2011: 226) suggests that archaeologists can use the concept of a least common denominator to compare terms: that data aggregators could combine together disparate specific terms by using a more generic term. There are some disadvantages to this method because broad common categories would obscure some of the details available in the original information. The least common denominator method proves most useful for information discovery, where a researcher could find all sites that fell within a certain broad time period even though all the sites were not directly contemporaneous.

Approaches to Data Integration

Organizations attempt data integration in different ways. The Digital Archaeological Record (tDAR), a digital repository run by Digital Antiquity, adapts Dublin Core metadata terms to assign generalized descriptions about each type of data stored (Digital Antiquity 2012). The Dublin Core Metadata Initiative, based in Singapore,

is a standards organization that defines general descriptive frameworks that researchers in different fields can customize to a very specific level for their fields (Dublin Core Metadata Initiative 2012). Researchers, by building a specific framework on top of the general Dublin Core framework, can make data encoded in a specific framework transferable back into the more general framework. For example, the tDAR category of "Site Name" could be encoded using the general Dublin Core category of "<dc:subject>." When integrating datasets, tDAR offers web-based data integration tools, which allow archaeologists to compare two or more tabular data sets (e.g., Excel spreadsheets, comma delimited text) and use the tools to personally combine different columns of data based on whether they believe them to be similar or dissimilar. The tDAR data integration tools allow researchers to exercise direct control and make conscious decisions about whether or not to combine different categories in datasets. This approach, however, demands a large investment of time and human effort to integrate datasets.

Other organizations are designing algorithms to integrate datasets automatically. OpenContext, the Web 2.0 dissemination tool created by the Alexandria Archive Institute, uses the ArchaeoML markup language to provide general categories for data (Kansa and Kansa 2011:60). ArchaeoML is a system of encoding "tags" that set off parts of a record file and describe the content of a piece of data (Schloen 2001:135). For example, ArchaeoML provides a tag of "<arch:property>" which denotes descriptive text like artifact type or method of creation. The tag is nested under ever more general category tags such as "<arch:properties>" and "<arch:observations>" which group together similar tagged data. In this example "<arch:properties> could also contain lines about artifact dimensions or excavation date. OpenContext adapts ArchaeoML system to experiment

with approaches to classifying data because the system allows both freeform keyword searches as well as highly structured hierarchical systems (Kansa and Kansa 2011:61). In the first case, artifacts may receive multiple time periods or cultural affiliations in order to reflect scholarly debate over ambiguity in terms. This approach is also human intensive, up front, but allows resources to be discovered using multiple terms and also better reflects when a resource is in dispute, rather than glossing over and codifying resources imprecisely.

The Integrated Archaeological Knowledge System (iAKS), developed by the MATRIX program at Michigan State University (MSU), utilizes a highly structured approach. As an information collecting and dissemination system, iAKS is focused on helping archaeologists collect and share new data rather than on accommodating existing data (Watrall 2011: 173). To this end, iAKS provides archaeologists choices of a number of controlled vocabularies for terms to apply to archaeological data. This highly structured approach, as Eiteljorg (2011:255) notes, controls, directs, and limits the kinds of archaeological inferences that can be made from the data and diverse datasets may be forced into pre-existing categories that do not make sense for the data. The approach offers a prime advantage, however, in that it allows archaeologists to easily combine multiple datasets created solely within the iAKS system.

Archaeologists outside the United States have also developed systems for archaeological metadata and vocabularies. The International Council of Museums provides a conceptual reference model (CRM), the CIDOC-CRM, for describing relationships and concepts in cultural heritage documentation (CIDOC Documentation Standards Working Group 2011) The base CIDOC-CRM contains a broad number of

categories for people, places, things, and times that define a hierarchy of terms and the relations between them. In the United Kingdom, employees of English Heritage and the University of Glamorgan Hypermedia Research Unit have created an extension to the CIDOC-CRM for archaeological terms called the CRM-EH (University of Glamorgan Hypermedia Research Unit 2012). These expanded terms allow archaeologists to more easily describe archaeological processes and concepts.

Conclusion

For many archaeologists, data integration represents one of the ultimate advantages of digital data: being able to aggregate small project data sets into a huge file that could be easily analyzed for widespread patterns. The reality of archaeological data, however, shows that data are often not compatible because of differing terms and a lack of strict standards. Strict standards are not realistic solutions either as archaeologists need to be able to customize their approaches to best fit their projects. Conceptual reference models, such as the CIDOC-CRM, ArchaeoML, or adaptations of Dublin Core provide a way for archaeologists to classify data at more general level, or provide multiple classifications to the same datum and allow archaeologists to integrate data. Archaeologists can adopt these models as ways to make their digital data easier for other archaeologists to discover. Approaches that accomplish that goal will improve data sharing in archaeology and may allow archaeologists to more easily study past human behavior on a larger scale.

Chapter 7
Who Owns Archaeology? Intellectual Property and Information Sharing

As mentioned in Chapter 3, the Archaeology Data Service is probably the most successful of the digital archaeological information repositories. This success is due in part to the different cultural resources laws that exist in the United Kingdom. Whereas in the United States, archaeological and other cultural resources are considered the property of the landowner on whose land they are found, in the United Kingdom, under the Treasure Act of 1996, items considered "treasure" can become the property of the government. Since many archaeological materials belong in the UK belong to the government, there is less need to make site locations confidential information. In the United States, because site looting and "pot hunting" are so rampant, laws exist to combat the loss of archaeological resources. In particular, the Archaeological Resources Protection Act of 1979 (ARPA) makes revealing, whether intentionally or through negligence, the location of an archaeological site illegal.

In addition to laws that govern archaeological information specifically, copyright laws sometimes pose barriers to sharing information. Copyright protections conflict with the spirit of the National Historic Preservation Act of 1966 (NHPA); the law underlying most archaeological work in the United States. The preamble of the NHPA states:

(1) the spirit and direction of the Nation are founded upon and reflected in its historic heritage;

(2) the historical and cultural foundations of the Nation should be preserved as a living part of our community life and development in order to give a sense of orientation to the American people;

If people –meaning all US citizens, not just archaeologists –do not have access to the information collected, how do they benefit and get a "sense of orientation" from cultural resources?

Copyright Law

US cultural resource management law provides a reason for cultural resource management archaeology to be done, as well as protects site locations by making them confidential information and not subject to Freedom of Information Act (FOIA) requests. The Freedom of Information Act is a US law that allows any person the right to obtain access to federal agency records, except in cases where exemptions apply – such as the provision in ARPA exempting archaeological site locations. The same regulations that protect site locations also make it difficult to share digital archaeological information.

US copyright law also creates obstacles to sharing archaeological information. The copyright for contract reports, such as those conducted by CRM firms on behalf of federal agencies and developers, is complicated. While reports commissioned by the federal government are supposed to be released to the public domain (17 U.S.C. § 105), unless properly otherwise specified in a CRM contract, the organization that paid for the cultural resource survey may hold copyright control over the report and other data collected during a survey or excavation. A major concern of cultural resource management archaeology is that most cultural resource management archaeology projects are considered "work for hire." "Work for hire" means an author voluntarily relinquishes his or her copyright and gives the copyright to the employer who commissioned the work (17 U.S.C. § 101). Unless explicitly stated in the contract, the company that funds

research may become the copyright holder for any report or project data sets (Nicholas and Bannister 2004: 332). In order to avoid this situation, federal agencies and cultural resource management firms need to explicitly state in their contracts that they intend to publish the results of the research and make clear who owns the copyright for a given project and related data set.

Copyright law typically only protects the implementation of an idea and not the idea itself (U.S. Copyright Office 2011:3). In addition, facts are not protected by US copyright (17 U.S.C. § 102(b)). Some degree of original creative input is required to copyright something. As such, data sets comprised of primary data are not eligible for copyright protections.

Alternatives to Traditional Copyright

US copyright uses an “all rights reserved” model that typically only allows reproduction with the express permission of the copyright holder. In the past decade however, the open-source software and the Creative Commons movements have arisen that utilize a "some rights reserved" model. Open source software licenses, such as the GNU General Public License (GPL) and Berkeley Software Distribution (BSD) licenses, allow users to freely reproduce computer code and documentation for noncommercial purposes (Engelfriet 2010). Different versions of the licenses may require users to also open source their code and documentation under similar licenses. Other versions of the licenses however, also allow for commercial use of licensed code. Creative Commons licenses operate in similar ways for non-computer code, such as written documents, music recordings, and video recordings. As with open-source software licenses, Creative

Commons licenses come in different forms: versions that allow for commercial use, noncommercial use, require attribution of the original author, and require that users licensing derivatives under the original license (Creative Commons 2012).

In addition to alternative copyright licenses, scientists and authors may also choose to commit their research to the public domain. According to the Creative Commons website (2012), data that researchers have explicitly removed copyright restrictions from and committed to the public domain allow other researchers the greatest flexibility to reuse scientific information. Wendrich (2011:227) emphasizes that the scientific culture of citing referenced sources would encourage attribution and credit for prior research. One advantage of public domain data is that the license negates the need to cite individual data sets in cases where thousands of data sets may be combined into one file and citing each of the individual sources may prove prohibitively difficult.

Cultural Concerns about Information Sharing

In addition to legal and copyright concerns about sharing data, there are also a number of cultural concerns. Indigenous groups, minority groups, and other historically disadvantaged populations have traditionally not had much say in what information museums, archaeologists, and anthropologists have collected, analyzed, written about, and shared about indigenous cultures (Nicholas and Bannister 2004: 339). Because of the lack of input, information has entered the public arena that groups wish had been kept confidential. This information includes pictures of human remains, photographs of burial artifacts, and detailed descriptions of sacred information such as religious ceremonies or the use of ceremonial paraphernalia (Nicholas and Bannister 2004: 335).

US copyright law does not protect traditional cultural intellectual property. US copyright law only protects "original works of authorship" that are fixed in a tangible form of expression (17 U.S.C. § 102 (a)). Because so much of traditional cultural intellectual property, say a story or a song, does not traditionally exist in a fixed form, copyright protections and rights are not conferred until someone puts the property into a fixed form, such as a video/audio recording or written document. Often, the first person to record a traditional cultural intellectual property is not a member of a traditional community, but is a non-native researcher, like an anthropologist or musicologist. Because of this, it is not unusual for traditional communities to not hold copyright to recordings of elders practicing in traditional ceremonies or telling traditional stories. Table 7.1 summarizes some differences between western and a small sample of views on intellectual property from held by some indigenous groups and explains why western laws are often insufficient to protect cultural intellectual property.

Table 7.1. Comparison of Indigenous and Western Views on Intellectual Property (adapted from Nakata et al. 2008).

Western Non-Indigenous Law	**Indigenous Custom**
Emphasis on material form – copyright only applies to "fixed" forms.	Generally orally transmitted.
Protections are limited in time (e.g., US copyright generally applies for 70 years after author's death; patent rights last for 20 years).	Emphasis is on preservation and maintenance of culture and typically protections have no set expiration period.
Individual creators emphasized – the "author"	Groups emphasized – property belongs/is entrusted to lineage groups.
Intellectual property is freely transmissible with author's permission –usually for economic gain – and can be converted to different forms.	Often intellectual property is not transferrable between groups. If groups allow transmission, the act must occur following a series of cultural qualifications.
Intellectual property rights holders decide how and to whom information is transmitted, transferred, and assigned.	Often groups have restrictions on how transmission can occur, especially with sacred or esoteric material.
Intellectual property rights are generally com-	A holistic approach, by which all aspects of

partmentalized into categories such as tangible, intangible, arts and cultural heritage.	cultural heritage are inter-related.
Emphasis on economic rights.	Emphasis on preservation and maintenance of culture.
No special protection of sacred secret material or gender restrictions.	Specific laws on gender and sacred information.

Table 7.1 lists very general concerns that are shared by many, but certainly not all indigenous communities. In order to properly protect culturally sensitive information, archaeologists must utilize their anthropological training to consult with indigenous communities about the concerns particular to each separate culture. A project to collect such concerns through a series of semi-structured interviews is outlined in Appendix A.

Culture of Science and Intellectual Property

Some archaeologists believe that they alone own the data that they collect (Eiteljorg 2002). If the goal of archaeology, however, is to build models that describe past human behavior, then archaeologists should be less concerned with who builds a model and more concerned with seeing that the model gets built. Archaeologists are only human, however, and often need be concerned with other things besides scientific advancement, such as getting tenure at a university, protecting their academic reputation, or simply feeling that since their collected data resulted from their hard work and the original collector of the data should be the person to benefit the most. Archaeology would benefit the most from an open sharing of archaeological data, however there are structures in place that provide powerful individual reasons for retaining information.

A major occupational concern of academic archaeologists is gaining tenure. Kansa and Kansa (2011: 69) note that traditionally, only peer-reviewed publications have been considered appropriate reason for a tenure committee to accept an application. Kansa and

Kansa (2011:71) however, suggest that incentives can be developed that encourage reluctant scientists to open up their data archives. By building a scientific culture that values data collection almost as much as interpretation, credit can be given for sharing primary datasets that rival the acknowledgement given for publishing research results.

Chapter 8
Guides to Good Practice: One "How-To" for Archaeological Information Archival

Since 1998, the Archaeology Data Service (ADS) has been publishing and updating "Guides to Good Practice" for digital archiving in archaeology (Archaeology Data Service and Digital Antiquity 2011a). Originally published as slim volume books, the Guides are now available online as webpages at http://guides.archaeologydataservice.ac.uk/g2gp/.

The ADS wrote the Guides to Good Practice to encourage archaeologists to better prepare data for long-term survival. Currently, many archaeological data are created in digital formats; whether reports written in word processing software, images taken using a digital camera, locational information recorded with a total station or GPS receiver, or more elaborate datasets created using three-dimensional laser scanners and other tools. A primary focus of the Guides is convincing archaeologists to preserve digital data in digital formats, rather than trying to convert information to a hardcopy format. Keeping digital data in a digital format retains the functionality of complex GIS, CAD, and relational database data while also preventing the need to redigitize data from hardcopies in the future. Besides retaining basic functionality in digital data files, archaeologists also may benefit from one of the primary advantages of digital data over paper records: the very low cost to reproduce and share a digital file.

In many cases, archaeologists expect to archive digital data in the same way that paper records are archived. Digital file formats change often, however, so simply preserving the media (e.g., compact disc, computer hard drive) is not adequate for truly preserving the archaeological data. When a computer program switches file formats, such

as Microsoft Word's transition from the various .DOC file formats to the .DOCX file extension, data encoded in the old file format often becomes unintelligible to modern versions of the program. Changes in computer hardware have a similar effect: rendering certain files and even certain storage media inoperable on newer hardware – after all, how many new computers are produced that still include floppy disk drives?

Project Lifecycles

In a perfect world, digital data go through several stages in a cycle of creation, use, and reuse. Researchers create data, describe and document their data through metadata, select which data to keep and which to discard, preserve and manage data, and then other researchers access and reuse the data. Professional archaeologists are already familiar with the first step: creating data. Some archaeologists may not be as familiar with the next step: describing and documenting data. Describing the form and characteristics of your data is critical for creating data that can be used by other researchers. Metadata, or "data about your data" includes descriptive information such as an identifier for the data, who created the data, when the data was created, the conditions of a recording event (e.g., the points per inch of a digital scan, weather encountered during an archaeological survey). Metadata may also include comments or helpful information such as the abbreviation codes used to speed up data entry (e.g., Lth_Mat, FGVR for Lithic Material, Fine-grained volcanic rock).

File Formats

Because file formats change so rapidly and often lose compatibility between

versions, the ADS are careful to recommend that archaeologists create and archive more stable archival formats along with their originally formatted datasets (Archaeology Data Service and Digital Antiquity 2011). Archival formats are usually files that are encoded in plain-text, using characters and symbols that appear as part of the basic American Standard Code for Information Interchange (ASCII) or Unicode character sets. The ASCII character set is based on the English alphabet, but also includes some symbols, while Unicode is much larger and can store information in many different languages. Even if a data format becomes unreadable by a machine, a human can still open an archival file in a text editor program and be able to understand much of the encoded data. For instance, spreadsheets or some database files could be encoded in a comma separated text file (CSV) where each column of information is separated by a comma and each new row is represented by a new line of data in the text file (Figure 8.1).

While not true archival formats, sometimes extremely popular proprietary file formats may be used as *de facto* standards for storing information. An example is the portable document format (PDF) from Adobe Systems. Before 1998, PDF was a proprietary file format whose structure was known only by Adobe Systems. Over time, Adobe worked to promote PDF as an International Standards Organization-recognized (ISO) format. The ISO accepted PDF 1.7 as a standard in 2008.

Documents

Text document formats are constantly evolving and frequently older versions of text document formats become incompatible with new software. The Guides note a movement of word processing utilizing more open standard file formats; such as the

XML-based .docx and .odt formats (Archaeology Data Service and Digital Antiquity 2011). Another large concern is that modern text documents often contain non-textual data such as images, tables, and sometimes even full spreadsheets and video. The Guides suggest a number of considerations when preparing to archive digital text documents. First, make sure that any content embedded within a document, such as images, spreadsheets, or videos are also archived separately and individually in the archival file formats appropriate for each content type. Separating the embedded content ensures that each file is accessible in the highest resolution or quality and is securely stored. Second, links within a document to external or changing content (e.g., a hyperlink to a website) will likely become non-functional over time, resulting in a "broken link." If the external links are important to the proper use of the main document, archival copies of the content they link to should also be made. Third, if an archaeologist chooses to preserve a document as a PDF file, they should also archive the version of the file used to create the PDF, such as the original Word or OpenOffice document. Archival formats for text documents include the .docx and .odt file formats. Both these formats comprise a zipped archive that separates the human-readable text, a file describing to the software how to format and style the text, and any additional content such as embedded images. Because the text is separated from the style and content files, the text may still be read in a simple-text editor, even if none of the other content is compatible with future software.

Databases and Spreadsheets

At first, databases and spreadsheets seem like very similar types of files. Structurally, however, the two are very different. Spreadsheets are modeled after paper

ledger sheets, and so are designed to be most useful in keeping track of and performing calculations on numerical transactions. Most spreadsheets are modeled after this paradigm. Databases are much more varied in their construction. The main database types are *flat files*, which follow a similar model as spreadsheets; *relational databases*, where data are split into many separate tables and linked together through a common category or common "key" value; and the growing *object oriented* model, where groups of variables form a digital object that reflects the characteristics of a real-world object or concept.

The Guides suggest that there are three main considerations when archiving databases and spreadsheets: embedded objects, data consistency and documentation, and non-data content. Embedded objects, as with text documents, are data in other file formats such as video or image files that are included within a spreadsheet or database. In order to ensure the best quality and easiest preservation, embedded objects are best archived as separate files. One of the main abilities of spreadsheets and databases are functions that can quickly summarize large datasets. The ability relies on all the data sharing common categories and being entered correctly, however. For example, if one field technician classified the color of a ceramic slip as "grey" and another field technician recorded the same slip as "gray," a database query for "Caram_Slip = 'grey'" would only return the first technician's records but ignore all of the second technician's records. Another concern is the use of codes for database and spreadsheet column headings and variable values. Without a sheet that explicitly lists translations of codes and abbreviations, an archaeologist would have a hard time making sense of a record written entirely using abbreviations. Non-data content refers to the various formatting

that can be applied to data in a spreadsheet or database such as font size and color, or cell style, color, and border thickness. While this content sometimes may superfluous, in many cases the formatting may indicate attributes about the data that would be lost in a conversion to a plain-text format, such as CSV.

Unlike most other file types, there is not a convenient, human readable format for images. Images come in two different kinds: raster images, where each picture element or "pixel" is represented by a tiny square with a certain value for color, hue, or other properties and vector images, in which the picture is represented by a series of geometric equations that describe the lines, curves, and fills of the picture. There are many different formats for each picture type, and important features to consider are whether the format compression loses fidelity (a "lossy" format), color depth, and transparency and metadata support. For raster images, uncompressed or lossless compression formats are best for preservation because they preserve the quality of the original image. Because of their small size and lossless compression, the Guides recommend portable network graphics (PNGs) for presentation (Archaeology Data Service and Digital Antiquity 2011). Tagged Interchange File Format files (TIFFs) are the recommended format for preservation, but their large file size makes them impractical for easy web dissemination. The Guides point out that JPEGs are particularly poor for preservation since they use lossy compression that degrades image quality. In addition, JPEG files pose a security risk since their decompression algorithm provides an exploitable flaw that allows malware to be installed in web browsers lacking the latest security updates.

Since vector images are stored as a series of geometric or vector equations rather than a matrix of pixels, computers can easily rescale vector images without a loss of

quality. In addition, vector images take up much less memory space than comparable raster images. Examples of vector images are Adobe Illustrator files, Scalable Vector Graphics (SVG files), and Portable Document Files (PDFs). Vector images also include many other files, but of the many types, the Guides only recommend SVG or PDF (specifically PDF/A; the subset of the PDF file designed for archival) files as suitable for preservation, since those file formats are based on open standards rather than closed, proprietary file format standards.

Video

Similar to images, there exists no "human-readable" format for video, only binary file formats that must be decoded by a computer to produce a visible moving image. The Guides suggest that the primary considerations for video files are ensuring the highest possible quality for master video files. The master video file should be of the highest possible quality and video resolution. The same master file should be used to produce dissemination copies. The dissemination copies can be edited down to only include a portion of the content present in the master file, and their resolution, framerate, and bitrate should be decided based on their anticipated use. Archaeologists should archive the master video file preferentially over the derived dissemination files since the dissemination files can be reproduced using the master video file, but the master video file can rarely be recreated using dissemination files.

Audio Files

Similar to video files, quality of files is the primary consideration for audio files.

Like video and image files, some audio formats are intrinsically of higher quality than others, either because they are uncompressed or because they utilize lossless compression. Although proprietary, waveform audio (WAV) files are suitable for preservation because they are commonly uncompressed and the format is widely used and openly documented. Other proprietary formats like Windows Media Audio (WMA) and RealAudio (RA) files are not suitable for preservation because the processes for creating the formats are protected by their owners and others may not reproduce the processes without breaking laws. Hence, if the format creator stops supporting the file format, no one else can recreate or decipher the file format. Lossy formats such as MP3 may be suitable for dissemination because of their small file size, but are not of high enough quality to be used for preservation.

GIS, CAD, and other Data Types

The Guides discuss more complicated data types such as shapefiles from geographic information systems (GIS), computer aided drafting (CAD), three-dimensional laser scans, virtual reality, and other data types. Because of the complexity of these topics, and because the Guides already present simplified summaries that necessarily leave out much information, readers are referred to the Guides to Good Practice chapters on these topics; and especially to each chapter's extensive bibliography.

Table format:

Column 1	Column 2	Column 3
A	B	C
1	2	3

CSV format:

```
Column 1,Column 2,Column 3

A,B,C

1,2,3
```

Figure 8.1. CSV provides a way to store tabular data as a string of comma separated values.

Chapter 9
HC SVNT DRACONES: The Unmapped Territories for Digital Data

When European explorers first began sailing across the Atlantic Ocean on voyages, maps of the world still contained large blank spots. In many of these maps cartographers used pictures of dragons to indicate the boundaries of the map, where the known world slipped into the unknown. Sometimes, near the dragon symbols cartographers wrote the Latin term "HC SVNT DRACONES" or "here are dragons." Fear of the unknown, represented by the dragons, kept many sailors close to shore and along the known, relatively safe routes. For a few brave explorers, however, the dragons presented a challenge. Through force of will, no small amount of luck, and many (sometimes disastrous) expeditions, the explorers pushed back the boundaries of the map and charted the unknown. In archaeology, there have been a few explorers into the digital realm (Eiteljorg II 2008; Schloen 2001; Richards et al. 2000; Kintigh and Altschul 2010) and more who have followed their lead (Kansa et al. 2011; McManamon and Kintigh 2010). Their explorations and experiments have proven valuable and shown directions where archaeologists may develop their approaches into viable workflows and tools for recording and using digital archaeological information. The explorers have also faced and experienced a number of pitfalls and mistakes along the way, such as the closing of the Archaeological Data Archive Project (Huggett 2006). The experiences here are not failures, but necessary steps for refining methods. From these pitfalls, archaeologists can now better differentiate between approaches that may work, and those destined to fail; a distinction that may not otherwise be intuitive.

How Archaeologists Can Integrate CRM and Academic Archaeology

Professional archaeology in the United States is split into two main groups: academic archaeology and cultural resource management (CRM) archaeology. While exceptions exist, academic archaeologists are mostly able to select the sites they choose to study and take the time to carefully analyze and interpret the data they collect. Academic archaeologists may spend years focused on only one or a few sites in order to fully analyze and interpret their collected data. Academic archaeologists rarely have the time to personally collect the broad datasets required to conduct far-reaching synthetic studies and so careful archaeologists must limit their interpretations of past human behavior to the specific areas and time periods for which they have the most detailed data. The goals of tenure, grant funds, and professional reputation encourage academic archaeologists to publish their interpretations, and sometimes their primary data, in peer-reviewed journals, edited volumes, professional meetings, and books.

Conversely, CRM archaeologists rarely select sites themselves and may only study sites that fall within the boundaries of a public land parcel or within the Area of Potential Effect (APE) of a project expected to adversely affect archaeological resources. In a best case scenario, CRM archaeologists will document a site's surface expression and successfully convince project proponents to change their project and avoid impacting the site. In these cases, the archaeologist will not excavate the site. Often sites are not avoided and so CRM archaeologists must focus on mitigating archaeological damage by excavating sites slated to be destroyed by construction projects. CRM archaeologists rarely have the time to fully analyze and interpret the data gained from salvage projects, but will prepare a report and listing of the data for limited publication. Often CRM

publications only reach the federal agency overseeing the project and the project proponent who paid for the study. Agencies do archive CRM publications, but these archives are much harder to access and search than published books and peer-reviewed journals. While some CRM companies and federal agencies do provide incentives for their archaeologists to analyze, present, and publish on collected data, many companies and agencies do not, which results in CRM archaeologists creating a backlog of collected data but rarely analyzing the data.

In short, CRM archaeologists possess a steady supply of opportunities and funding for archaeological fieldwork and data collection. Academic archaeologists possess resources and incentives to carefully analyze data and professionally publish their interpretations. Wider publication and easier access to broad-scale CRM data sets would enable academic archaeologists to interpret past human behavior on much larger scales than currently possible and would encourage the proper use of the data CRM archaeologists collect. Digital data archival and dissemination offers the missing factor that would allow CRM and academic archaeologists to enter into a relationship that not only benefits archaeologists personally, but also improves the quality of archaeology as a discipline.

Importance of Preserving Digital Data

Digital formats offer archaeologists possibilities for recording, analysis, and dissemination of archaeological data that paper records simply cannot match. Electronic records provide cheap reproduction and transport of huge reports and datasets. Archaeologists may also more easily combine electronic datasets to create aggregate

datasets representing fieldwork across wide study areas. Some extremely useful electronic records, such as computer aided drafting (CAD) drawings, three-dimensional laser scans, and geographic information system (GIS) datasets lose utility and do not translate well into hardcopy formats.

With easier reproduction and transport, more archaeologists can access datasets and test past interpretations using new methods and theories. From a processual archaeology perspective, retesting and analyzing data allows archaeologists to better support and refine interpretations and furthers scientific validity. From a post-processual perspective, cheaper and easier access allows more people to interpret archaeological data, including populations who archaeologists have historically distanced, such as Native Americans.

Different archaeological theories explain why archaeological data, like archaeological sites, are non-renewable and should be preserved.

Combining Digital Data

Digital formats allow archaeologists to combine individual datasets from across a broad study area into one large aggregate dataset. Aggregate datasets allow archaeologists to make broader interpretations of past human behavior and allow archaeologists to compare human behavior in one area with human behavior at others. Pulling together data from across a broad study area allows archaeologists the most confidence in determining whether a behavioral pattern actually exists or whether single sites represent anomalies.

Sharing Digital Data Responsibly

Digital formats make sharing archaeological information easy, but sometimes there are valid reasons for not sharing information. Federal laws restrict the sharing of archaeological site locations in order to prevent damage and complicated copyright laws make it difficult to determine who has the right to share a piece of information. Archaeological collections sometimes contain information that Native American tribes find sacred and sharing such information with people from the wrong clan, the wrong gender, or non-Native Americans would reduce the power of related ceremonies. Archaeologists need to consider all these factors when deciding whether a digital file should be shared.

Conclusion

As archaeologists continue to work in the modern field, they face a number of challenges stemming from increasing information flow and overload, demands to demonstrate archaeology as a worthwhile endeavor for humankind, a need to justify the use of public money, and decreased funding for traditional excavation-focused archaeology. Digital tools allow archaeologists to meet and exceed the demands politicians, diminishing research funds, and archaeological publics place on archaeologists, but only if archaeologists learn to use the tools properly. Archaeologists need to integrate digital tools into their planning and thinking processes, learn how to use tool at least at a basic level, and think critically about how to control for biases and limitations imposed by digital tools. Archaeologists need to rethink scholarly publication

and how to respect intellectual property, both legally and culturally. The great benefits of better preservation, curation, and dissemination standards are numerous, and worth the great effort required.

Open-access to a wider symbolic representation of the archaeological record, reanalyzed periodically by archaeologists using various new, old, and yet to be created theoretical positions and by comparison with the continually expanding body of archaeological evidence can account for and limit the influence of interpretive biases. Wide dissemination of archaeological results and datasets offers diverse and often disenfranchised publics a voice in interpretations and heritage and also lets people learn about real archaeological practice. Well-defined, yet evolving, standards of planning, organization, preservation, interoperability, and open access of digital databases represent a pathway towards creating, preserving, and disseminating a digital archaeological record.

References Cited

Archaeology Data Service, and Digital Antiquity
2011 Guides to Good Practice. York, England / Tempe, Arizona: Archaeology Data Service / Digital Antiquity. http://guides.archaeologydataservice.ac.uk/, accessed February 8, 2012.

Binford, Lewis R.
1965 Archaeological Systematics and the Study of Cultural Process. American Antiquity 31(2): 203–210.

1968 Some Comments on Historical Versus Processual Archaeology. Southwestern Journal of Anthropology 24: 267–275.

1980 Willow Smoke and Dogs' Tails: Hunter-Gatherer Settlement Systems and Archaeological Site Formation. American Antiquity 45(1): 4–20.

CIDOC Documentation Standards Working Group
2011 The CIDOC CRM. CIDOC Documentation Standards Working Group. http://www.cidoc-crm.org/, accessed February 14, 2012.

Creative Commons
2012 About. Electronic document, http://creativecommons.org/about, accessed March 5th, 2012.

Digital Antiquity
2012 Digital Antiquity, http://www.digitalantiquity.org/, accessed March 5th, 2012.

Digital Antiquity
2012 Professional Advisory Panel, http://www.digitalantiquity.org/about/scienceadvisors/ accessed March 5th, 2012.

Dublin Core Metadata Initiative
2012 DCMI About Us. Dublin Core Metadata Initiative Limited. http://dublincore.org/about-us/, accessed February 14, 2012.

Eiteljorg II, Harrison
2002 The Archaeological Data Archive Project. Center for the Study of Architecture. http://csanet.org/archive/adap/, accessed February 3, 2012.

2004 Computing for Archaeologists. In A Companion to Digital Humanities P. Chapter 2. Oxford: Blackford. http://www.digitalhumanities.org/companion/, accessed January 5, 2012.

2008 VIII Protecting and Preserving the Archaeological Record. In Archaeological Computing. 2nd edition Pp. 241–252. Bryn Mawr, PA: Center for the Study of

Architecture.

2011 What Are Our Critical Data Preservation Needs? In Archaeology 2.0: New Approaches to Communication & Collaboration. Eric C. Kansa, Ethan Watrall, and Sarah Whitcher Kansa, eds. Pp. 251–264. Cotsen Digital Archaeology Series. UC Los Angeles: Cotsen Institute of Archaeology http://escholarship.org/uc/item/1r6137tb.

Elston, R.G.
1992 Archaeological Research in the Context of Cultural Resource Management: Pushing Back in the 1990s. Journal of California and Great Basin Anthropology 14(1): 37–48.

Engelfriet, A.
2010 Choosing an Open Source License. Software, IEEE 27(1): 48 –49.

Fuka, Lauren
2009 Digital Repatriation and Indigenous Communities: Connecting Museum Collections and Source Communities in Australia and the Pacific. Anthropology, Northern Arizona University Anthropology Library, Northern Arizona University.

Green, William, and John Doershuk
1998 Cultural Resource Management and American Archaeology. Journal of Archaeological Research 6(2): 121–167.

Hodder, Ian
1996 Theory and Practice in Archaeology. Reprinted. New York: Routledge.

Huggett, Jeremy
2006 Ten Years of the ADS. Archaeology Data Service. http://ads.ahds.ac.uk/newsletter/issue19/10yrs.html.

ICOMOS Committee on Archaeological Heritage Management
2002 Heritage at Risk 2001-2002: Archaeological Sites at Risk. International Council on Monuments and Sites. Electronic document, http://www.international.icomos.org/risk/2001/icahm2001.htm, accessed May 5, 2012.

Kansa, Eric C., and Sarah Whitcher Kansa
2011 Toward a Do-it-yourself Cyberinfrastructure: Open Data, Incentives, and Reducing Costs and Complexities of Data Sharing. In Archaeology 2.0: New Approaches to Communication & Collaboration. Eric C. Kansa, Sarah Whitcher Kansa, and Ethan Watrall, eds. Pp. 57–92. Cotsen Digital Archaeology Series. UC Los Angeles: Cotsen Institute of Archaeology http://escholarship.org/uc/item/1r6137tb.

Kansa, Eric C., Sarah Whitcher Kansa, and Ethan Watrall, eds.
2011 Archaeology 2.0: New Approaches to Communication & Collaboration. Cotsen Digital Archaeology Series. UC Los Angeles: Cotsen Institute of Archaeology. http://escholarship.org/uc/item/1r6137tb.

King, Thomas F.
2004 Cultural Resource Laws and Practice : an Introductory Guide. 2nd edition. Walnut Creek, CA: AltaMira Press.

Kintigh, K. W, and J. H Altschul
2010 Sustaining the Digital Archaeological Record.

Limp, W. Frederick
2011 Web 2.0 and Beyond, or On the Web, Nobody Knows You're an Archaeologist. In Archaeology 2.0: New Approaches to Communication & Collaboration. Eric C. Kansa, Ethan Watrall, and Sarah Whitcher Kansa, eds. Pp. 265–280. Cotsen Digital Archaeology Series. UC Los Angeles: Cotsen Institute of Archaeology. http://escholarship.org/uc/item/1r6137tb.

McManamon, F. P, and K. W Kintigh
2010 Digital Antiquity and the Digital Archaeological Record (tDAR): Broadening Access and Ensuring Long-Term Preservation for Digital Archaeological Data.

McManamon, Francis P.
2011"Digital Antiquity: Enabling and Enhancing Preservation of and Access to Archaeological Information" Grant: 2010-2011 Report. Digital Antiquity.

Moore, G.E., and others
1998 Cramming More Components onto Integrated Circuits. Proceedings of the IEEE 86(1): 82–85.

Nakata, Martin, Vicky Nakata, Alex Byrne, et al.
2008 Australian Indigenous Collections: First Generation Issues.

Nicholas, George P., and Kelly P. Bannister
2004 Copyrighting the Past? Emerging Intellectual Property Rights Issues in Archaeology. Current Anthropology 45(3): 327–350.

Oppel, Andy
2004 Databases Demystified. McGraw-Hill, Emeryville, CA.

Presland, G.
1984 A Cautionary Tale: Computerization of the Victoria Archaeological Survey Sites Register. Advances in Computer Archaeology 1: 7-11.

Reilly, Paul and Sebastian Rahtz

1992 Introduction: Archaeology and the Information Age. In Archaeology and the Information Age: A Global Perspective, edited by Paul Reilly and Sebastian Rahtz, pp. 1-21. Routledge, London.

Richards, Julian D. and N.S. Ryan
1985 Data Processing in Archaeology. Cambridge University Press, Cambridge.

Richards, Julian, and Archaeology Data Service.;Arts and Humanities Data Service.
2000 Digital Archives from Excavation and Fieldwork : a Guide to Good Practice. 2nd ed. Oxford;Oakville CT: Oxbow ;;David Brown Book Co. [distributor].

Richards, Julian, Stuart Jeffrey, Stewart Waller, et al.
2011 The Archaeology Data Service and the Archaeotools Project: Faceted Classification and Natural Language Processing. In Archaeology 2.0: New Approaches to Communication & Collaboration. Eric C. Kansa, Sarah Whitcher Kansa, and Ethan Watrall, eds. Pp. 31–56. Cotsen Digital Archaeology Series. UC Los Angeles: Cotsen Institute of Archaeology http://escholarship.org/uc/item/1r6137tb.

Schloen, J. David
2001 Archaeological Data Models and Web Publication Using XML. Computers and the Humanities 35(2): 123–152.

Schöpfel, Joachim
2011 Towards a Prague Definition of Grey Literature. In Twelfth International Conference on Grey Literature: Transparency in Grey Literature. Grey Tech Approaches to High Tech Issues. Prague, 6-7 December 2010 Pp. 11–26. TextRelease. http://archivesic.ccsd.cnrs.fr/sic_00581570/fr/.

Seymour, Deni
2010 Sanctioned Inequity and Accessibility Issues in the Grey Literature in the United States. Archaeologies 6(2): 233–269.

Shanks, M., and C. Tilley
1987 Re-Constructing Archaeology. Cambridge University Press, Cambridge.

Schiffer, Michael B.
1972 Archaeological Context and Systemic Context. American Antiquity 37(2): 156-165.

Trosper, Ronald
2002 Northwest Coast Indigenous Institutions that Supported Resilience and Sustainability. Ecological Economics 41: 329-344.

University of Glamorgan Hypermedia Research Unit
CRM-EH: English Heritage Extension to CRM for the Archaeology Domain.

University of Glamorgan. http://hypermedia.research.glam.ac.uk/kos/CRM/, accessed February 14, 2012.

Walker, Phillip L., Clark Spencer Larsen, and Joseph F. Powell
2000 Final Report on the Physical Examination and Taphonomic Assessment of the Kennewick Human Remains (CENWW.97.Kennewick). In Report on the DNA Testing Results of the Kennewick Human Remains from Columbia Park, Kennewick, Washington. National Park Service http://www.nps.gov/archeology/kennewick/walker.htm.

Watrall, Ethan
2011 iAKS: A Web 2.0 Archaeological Knowledge Management System. In Archaeology 2.0: New Approaches to Communication & Collaboration. Eric C. Kansa, Ethan Watrall, and Sarah Whitcher Kansa, eds. Pp. 171–184. Cotsen Digital Archaeology Series. UC Los Angeles: Cotsen Institute of Archaeology. http://escholarship.org/uc/item/1r6137tb.

Wendrich, Willeke
2011 UCLA Encyclopedia of Egyptology, Archaeological Data, and Web 2.0. In Archaeology 2.0: New Approaches to Communication & Collaboration. Eric C. Kansa, Sarah Whitcher Kansa, and Ethan Watrall, eds. Pp. 211–232. Cotsen Digital Archaeology Series. UC Los Angeles: Cotsen Institute of Archaeology. http://escholarship.org/uc/item/1r6137tb.

Appendices

Appendix A

Interviews to Identify Cultural, Professional, and Educational Needs
for Digital Archaeological Repositories
(CPENDAR)

Russell Alleen-Willems
Research Design

Digital Archaeological Repositories

Archaeologists have long been aware of the value of archaeological repositories for physical remains and associated records. Archaeologists understand that archaeology is a science whose process destroys the object of study. Because of this fact, archaeologists must protect and preserve individual artifacts, samples, detailed descriptive notes and other associated data, such as measurements, as these items are integral to allowing confirmation or denial of archaeological interpretations by reanalysis (McManamon and Kintigh 2010). S. Terry Childs (2008) has conducted research about archaeological repositories since the early 1980's and notes that there is a "curation crisis" in archaeology. When faced with funding shortages, managers and administrative officials too often see archives as easy places to save money by cutting budgets. Childs focuses mostly on physical archives. However, archives of digital archaeological are in danger as well. Today, archaeologists create and store much archaeological data primarily or entirely in digital formats (Kintigh and Altschul 2010).

Archaeological reports created for compliance with Section 106 of the National Historic Preservation Act of 1966 typically receive only limited publication in hardcopy but are usually created as computer word processing documents or portable document format files (PDF's). Reproducing and distributing these digital files costs substantially less than creating hard copies and often archaeologists trade these documents informally on an as-needed basis. This informal method works but has one large problem: archaeologists find it exceedingly difficult to fully inventory the previous archaeological research conducted in an area. To compile even a partial list of prior work, archaeologists need to consult disparate resources such as local Bureau of Land Management offices,

State Historic Preservation Offices, National Forest Service and National Park Service offices. Even after consulting these resources, wise archaeologists also contact other archaeologists in their professional network to see if they have conducted or heard of any other projects. The picture of prior archaeological work compiled in such a manner is likely not comprehensive and, as a result, archaeologists sometimes repeat work unnecessarily (McManamon and Kintigh 2010).

In addition, archaeologists typically store primary data (e.g., measurements, maps, stratigraphic and structural drawings) in a number of different file formats. These files are usually not included directly in archaeological reports, but may be used as the basis for creating figures and tables for archaeological publications. These publications may only present a portion of the entire data collected and archaeologists typically focus on the subsection of data pertinent to specific research questions. A recent trend is for archaeologists to publish even less of their primary data in hard copy, using "thin volumes" that disseminate the results of research but link to online versions of primary data appendices (Archaeology Data Service 2011).

Organizations in the United States and around the world have begun digital repository projects that seek both to preserve archaeological information in digital formats and to use the cheap cost of reproducing digital files to provide wider access to archaeological information (McManamon et al. 2010). Most of the research these organizations conduct focus on developing methods for creating and storing information in digital file formats that will remain accessible and readable after the hardware used to create files has become obsolete. This research also centers on deciding what fee repositories should charge to preserve digital data in perpetuity (Archaeology Data

Service 2011). These organizations have focused on the necessities of creating and funding archaeological repositories, but have yet to research much about what potential end users need archaeological repositories to do. Different users may have different needs, whether in terms of requiring data in specific formats, protecting legally or culturally sensitive data, ensuring that legal and cultural intellectual property rights are respected and especially to what uses non-archaeologists would put archaeological data.

Study Purpose

The purpose of this study is to collect critiques, concerns and suggestions to improve digital archaeological repositories. I will use the responses from interviews to construct questions for a structured survey that can be administered to a wider population than can be interviewed in person. Repository creators can use the concerns and suggestions identified to design digital archaeological repositories that meet the needs of users with different interests and backgrounds while also protecting site confidentiality and addressing issues of intellectual property.

Study Populations

I will solicit interviews from individuals representing several different communities. I plan to target archaeologists in the Federal, Cultural Resource Management, Native American and academic sectors; educators in K-12 schools; museum educators; and participants in avocational archaeology and historic interest groups. At least some of these groups will require Institutional Review Board (IRB) approval to interview. I will interview people from these groups if Northern Arizona University and tribal governments approve my application.

I will select archaeologists and other participants to interview by using my network of professional contacts. I will ask my network to refer interested participants for me to interview and I will interview people I personally know. I will use a similar process to gather participants who teach in K-12 schools. I will utilize a list of museums who responded to a survey about digital repository pricing to identify and solicit participants for interviews that represent museum educators (Watts 2011). Finally, I will send out calls for participants on mailing lists and online forums for avocational archaeology groups, as well as directly contacting leadership of historical societies.

Study Methods

Qualitative cultural expert interviews form the methods for this research. I will use cultural expert interviews to identify topics around which to write questions for an online survey, which I may send to a wider audience than I could interview on an individual basis.

I will conduct interviews over the course of the Fall 2011 and Spring 2012 semesters at Northern Arizona University (August 29, 2011 – May 11, 2012). For the sake of completing the research project in a reasonable amount of time, I intend to interview twelve participants, although I will consider interviewing more if time and the number of participants allow. This sample should help me tailor my survey questions to solicit relevant information. I will conduct as many interviews in person as possible, but will also utilize phone calls and the video conferencing program "Skype" to collect interviews from participants outside of my immediate area. I will record interview audio using a digital voice recorder for in-person interviews and software plug-ins for

interviews conducted online or via the phone. Along with detailed notes and a final version of my report, I will store these audio recordings on a personal external hard drive that I will backup to a DVD disc once the research project is complete. The file names of the audio files will contain the date of the interview, the interview group and an alias for the participant but will not reference the real identity of the participant in any way (e.g., 20110717_k12teach_Teach1.mp3). The DVD will remain in my personal archive after research completion.

I will analyze the data by listening to the interviews to identify key concepts relating to my research questions, as well as any other dominant themes. Once I identify key concepts, I will segment the interviews into discrete units that I can analyze and code accordingly. I will quote specific sections of these interviews to give examples of typical responses to each of the research questions, as well as examples of themes I did not think to include in this design.

I will report on the interview results in a chapter of my master's thesis on digital archaeological repositories. This chapter will also include the survey questions I create, which I may use to write a grant for future research into the topic of user needs for digital archaeological repositories.

Research Questions

I designed this research project seeking to answer the following research questions:

1. What, if any, kinds of archaeological information should be shared through the internet?
2. Who should be able to access archaeological information?

3. Who should make the decision about what archaeological information to share?

4. If archaeological information were available online, what would you want to be able to do with it?

I based the first research question on my hypothesis that members of different groups will have different ideas on what kinds of information are appropriate to share with different audiences. For example, I expect that Native American individuals may want to restrict some archaeological information that needs to be secret to members of a specific clan, or that relates to ceremonies that diminish in power if people reveal the ceremonies to the wrong parties. I also expect that many, but possibly not all, professional archaeologists would want to restrict site location information in order to protect sites from destruction and vandalism.

The second research question relates closely to the first but approaches the topic from a different angle. I hypothesize that different groups think that archaeological information can be appropriately shared through the internet but wish to restrict that sharing only to certain groups or individuals. For example, archaeologists working in federal service may wish to share information only with other federal archaeologists. Similarly, Native American individuals may want to share information only with other Native Americans.

The third question relates to concerns I gathered through informal talks with other archaeologists. Online archaeological repositories offer access to reports from different United States National Parks and Forests. Archaeologists I spoke to work at some of these National Parks and related that lead archaeologists for different Parks and Forests

were dismayed and angry that someone uploaded archaeological information about "their" areas without permission. While repository creators need to consider the question of intellectual property and copyright law, I am interested in how different people view archaeological information as resources that can be owned and controlled by specific groups or individuals.

I designed the fourth research question to elicit responses that detail the wide number of uses for archaeological information, if available online and to a wide body of users. I hypothesize that repository creators have not anticipated many of these uses but that these uses could benefit both archaeologists as well as the many publics archaeologists should serve.

Cultural Expert Interview Questions

I will conduct the interviews using the following set of questions. I will always ask the numbered questions and may ask the lettered questions to follow up on conversation topics or as probes to solicit deeper responses or clarification. I will also use the probe question "why?" to get at reasons behind responses that I may not immediately understand.

1. What do you think are responsible ways to share reports about archaeological surveys, excavations or analyses?

a. What, if any, are irresponsible ways to share reports?

2. What do you think are responsible ways to share archaeological information, such as primary data like measurements or radiometric dates?

a. What, if any, are irresponsible ways to share primary data?

3. What, if any, kinds of archaeological information should be shared through the internet?

a. If the internet is an appropriate method for sharing information, what, if any, kinds of information should be protected?

b. What ways would you suggest protecting information?

4. Who should be able to access archaeological information?

a. Should access of archaeological information of different kinds be restricted to specific groups?

4.a.i. How would you restrict archaeological information?

b. What archaeological information is appropriate for different audiences, such as an 8th grade history class instead of an archaeologist at a university or an archaeologist working for a museum instead of a member of a Native American group?

5. Who should make the decision about what archaeological information to share?

a. What gives someone the right to control a piece of archaeological information?

b. Who should be involved in the decision to share a piece of archaeological information?

6. If archaeological information were available online, what would you want to be able to do with it?

a. In what file formats would you need archaeological reports to be available for your uses?

b. In what file formats would you need archaeological primary data to be available for your uses?

c. For what purposes do you think a museum could use archaeological information?

d. For what purposes do you think a K-12 educator could use archaeological information?

e. For what purposes do you think a Native American group could use archaeological information?

Uses of this Research

I seek to improve digital archaeological repositories by collecting critiques, concerns, suggestions and requests from the population of people who use, or might use, digital archaeological repositories. I believe that digital archaeological repositories could benefit a wide number of people, but only if repository creators are aware of the different needs that user groups possess. I will collect this information using qualitative cultural expert interviews that I will analyze to identify key concepts that repeat among many groups, as well as concerns and needs specific to individual groups. I hope that repository creators will acknowledge and seek to meet the needs I identify through this study for both the benefit of archaeological repositories and for the many publics archaeologists should serve.

References Cited

Childs, S. Terry, and Seth Kagan
2008 A Decade of Study into Repository Fees for Archaeological Collections. Studies in Archaeology and Ethnography. Number 6. Archaeology Program, National Park Service, Washington D.C.

Kintigh, Keith W., and Jeffrey H. Altschul
2010 Sustaining the Digital Archaeological Record. Heritage Management. 3(2):264-274.

McManamon, Francis P. and Kintigh W. Kintigh
2010 Digital Antiquity: Transforming Archaeological Data into Knowledge. SAA Archaeological Record. March 2010, p.37-40.

McManamon, Francis P., Keith W. Kintigh and Adam Brin
2010 Digital Antiquity and the Digital Archaeological Record (tDAR): Broadening Access and Ensuring Long-Term Preservation for Digital Archaeological Data. Newsletter of the Center for the Study of Architecture. Vol. XXIII, No. 2.

Archaeology Data Service
2011 Planning for the Creation of Digital Data. In *Guides to Good Practice*. Edited by Kieron Niven. Archaeology Data Service, University of York, United Kingdom.

Watts, Josh
2011 Policies, Preservation, and Access to Digital Resources: The Digital Antiquity 2010 National Repositories Survey. Manuscript on file, Digital Antiquity, Tempe, Arizona.

Call for Participation

If you are interested in the use of archaeological information in a digital format, I would like to talk to you.

I am Russell Alleen-Willems, an Applied Archaeology graduate student at Northern Arizona University. I am conducting a study called "Interviews to Identify Cultural, Professional, and Educational Needs for Digital Archaeological Repositories" (CPENDAR). Through one-on-one interviews, this study collects comments, concerns, criticisms, and suggestions for the design and use of digital archaeological repositories. The CPENDAR study will form part of my Master's thesis, and interview results will be used to form questions for a structured survey that collects similar information from the wider body of interested archaeologists, educators, and interested publics. I am particularly interested in how participants view topics such as how to best protect legally or culturally sensitive information and who should have access to, or control over, archaeological information.

I am currently scheduling in-person, phone-based, and Skype-based interviews. Each interview will last approximately one hour and will be recorded using digital voice recorder or audio recording software. All information you provide will be anonymized in any publications and your participation will be kept confidential.

Please email me at ralleenwillems@gmail.com if you are willing to be interviewed.

Thank you in advance for your time. If you have any questions or concerns please do not hesitate to contact me.

Sincerely,

Russell Alleen-Willems
Graduate Student, Applied Archaeology
Anthropology Department
Northern Arizona University
ralleenwillems@gmail.com
360-441-9416