

Digitization and the Murray State University Archaeology Program

Kit W. Wesler
Murray State University

Federal agencies are promoting the idea that public sponsorship of research produces records and collections held in public trust, which therefore should be open to public scrutiny. Making data available on the Internet is one way to fulfill this requirement. Computer technology allows for powerful methods of data storage, analysis, presentation, and publication of archaeological data. The Murray State University Archaeology Laboratory is digitizing and making accessible all field and laboratory data, as well as reports, for collections and research that we curate. We have embarked on re-cataloguing our full collection, designing digital versions of all field record forms, and adding an Archaeological Information Systems emphasis to our M.Sc. program in Geosciences.

A paper for presentation to the Twenty-sixth annual Kentucky Heritage Council Archaeological Conference, Murray, Kentucky, March 6-8, 2009.

The Murray State Archaeology Lab is now just over 30 years old. Although there had been a couple of archaeologists working at Murray in the 1970s, they did not leave a systematic organizational structure behind. When Ken Carstens arrived in the fall of 1978, he found a bunch of stuff lying around, including bags of artifacts with no useful provenience. He recognized the need for an accessioning system, which he created.

Now, some of us remember 1978. Portable computers were barely even seen on *Star Trek*, let alone on office desks. The original lab accession system was adequate for the time, based on a year and serial number designation for each collection that came in. It was hard to look thirty years ahead and realize that, between the sheer size of the collection and the opportunities and needs of computer cataloguing, database design and analysis, complexity would increase fast. I won't even mention the challenges of expanding space and curation requirements.

Computers have generated opportunities unpredicted in 1978, but also are creating expectations. We are already hearing calls for open access to our data. In 2006, Senators John Cornyn (R-TX) and Joe Lieberman (D-CT) introduced the Federal Research Public Access Act. The Act would have required that eleven Federal agencies that fund the bulk of government-sponsored research make all manuscripts for publications resulting from sponsored research publically accessible by posting them on the Internet, and hosting a web archive to maintain open access to them (U.S. Congress 2006; Kansa 2007:1).

The bill didn't pass, but it's expected to be reintroduced. It's also expected to be the leading tip of a new wave of expectations for open access to data. The organizers of the SAA's Digital Data Interest group expect that, eventually, all research funded by or permitted by

Federal action will be required to be accessible to the public. This includes by far the majority of all archaeology conducted in the United States.

The reasoning behind this is fairly straightforward, and fits quite well with our own stated ethic of stewardship. Archaeology is supported as a public trust, because our Federal and state governments value the preservation and investigation of our common heritage. Archaeology is vested in four major types of institutions: universities, museums, government land-management, planning, granting and permitting bureaus, and private cultural resource management firms. All of these institutions derive much of their funding from public sources, even private CRM firms whose work is mandated by law, regulation or permit action. Archaeological ethics emphasize stewardship, in the preservation of sites and collections from sites, the publication of reports, and, increasingly, the further dissemination of results in forms accessible to the general public (which we call “public archaeology”). We recognize that the collections that we manage in laboratories or museums include artifacts and also the systematic information that places the artifacts in context. The field and laboratory records that document the artifacts give them meaning.

As Keith Kintigh (2006:572) notes, the ethical principles of archaeology should promote the preservation of and access to our raw data as well as interpretive publications. However, the raw records are generally not accessible to the public. They are available only to those fellow researchers who make the journey to the archives and gain physical access to them. This situation leaves the full process of archaeology in the realm of mystery as far as the general public is concerned. The public remains convinced that either the conduct of archaeology is arcane, esoteric and perhaps even deliberately hidden from view in some vast conspiracy that feeds the paranoiac (but lucrative to publishers) fantasies of pseudoarchaeology, or that it is little different from treasure-hunting, since only the finest artifacts and the most entertaining stories make it to the exhibit and the television screen.

Archaeologists have always promoted the idea of publication. But publications are limited. Most—as in any academic discipline—are addressed to colleagues in a technical literature. Some publications attempt to speak to the more general audience, with varying success. None attempts to publish all the data, or all of the field documentation, for several reasons. For one, many archaeologists see the raw data as proprietary, even when supported by public funds. This view may relate to a sense that the analysis is never truly finished, but that the researcher intends to finish it some day (and will not share credit with potential follow-up researchers). Or: much of the raw data is seen as too technical to be understood by a lay reader, and perhaps too messy (or maybe too slipshod?) to acknowledge. Then again, the cost of publishing all of the notes, data tables, drawings and photos has simply been prohibitive in a hard copy format.

The first two objections can be countered with a simple formula: public funding of research results in records held in public trust, and therefore should be open to public scrutiny (with some safeguards for site location and ownership data, to protect sites and privacy).

The third objection is becoming less formidable as publishing technology has evolved. We’ve already seen publications on CD-ROM, hybrid print-and-CD publications, and Internet

dissemination of reports, as well as internet-accessible queryable databases like the Digital Archive of Comparative Slavery hosted by Monticello (DAACS 2004).

The problem for us—speaking here specifically for the Murray State Archaeology Lab—is how to get ready for the brave new world in which we will be expected to make our data available. I've been thinking about this for a couple of years now, and I'd like to describe how we're going about it. I do NOT offer my system as the model that others should follow. But I hope that I can further a general conversation about ways by which we can move forward to create accessible archives.

Like most of us, I use a variety of recording forms in the field and lab. In the last couple of years, I have developed digital versions of all field recording forms used in his research. These records include: a Daily Log, a Square Sheet (the excavation unit map; this form has a reverse with data on individually-mapped artifacts), a Profile Elevation form, an Elevation record (which I just keep on a pad), a Shovel/Posthole Test data form, a Field Inventory, and a Photo Data record. I'm replacing a formerly-used Feature Data form with a newly-designed Context form reflecting the needs of Harris-style recording (Harris 1989), which has become standard in Europe but has yet to be adopted widely in the United States. The digital Context form, Square Sheet, and Profile Elevation form incorporate hyperlinks to field drawings and photos in Windows bitmap or other graphic formats.

The database files, with digital entry forms, are in Microsoft Access. They carry the same information as my paper-based forms, in roughly the same layout. I can hyperlink to the field drawing, in legacy cases scanned from a paper form. With a tablet style laptop, these graphics can be drawn directly on the computer. I also hyperlink to a photo. How the graphics come up depends on the default settings on the individual computer. My goal is to be able both to incorporate previous field notes by data entry and scanning, and to create these records directly in the field without paper.

Additional digital files that form the corpus of documentation for the project are the artifact catalogue, field and periodic reports, and analytical databases.

All of these digital formats use standard software, mostly compatible with web browsing and/or download. Text files created in MS Word and Excel will be printed to Adobe.pdf format for posting on the web.

The next step, of course, is posting them to the web. I'm not an html programmer and don't really want to be. So I have a new homepage for the MSU Archaeology Lab: <http://infosys.murraystate.edu/KWesler/default.aspx>, which uses MS Windows Share Point Services 2.0, a file-sharing software, as the platform and portal. The home page will contain announcements (for instance, updates), links to each project archive, and links to supporting documentation such as the MSU Archaeology Lab procedures manual and collections policy. I also have a research tools section for handy items that I use or have developed for my projects, that may be of interest to my students.

I had intended to make this publically available to coincide with today's paper as a sort of grand announcement, but by accident it actually went up about a month ago.

This is not by any means a perfect system. There are problems making some of these files accessible. The main problem is MS Access. Access files are treated as executable programs by most security systems, meaning that visitors to the site cannot either open my databases remotely or download them to open on their home computer. We can't email Access files, either. I have posted them here as html tables and also as zipped files. If a visitor wants to use the hyperlinks, he or she will have to download not only the database but also the graphics folders. But at least the basic data are here.

The solution to this is to design relational databases and publish them to the 'net through a SQL server. We don't yet have that capability, but we will be working toward it. In the meantime, I can at least get a lot of data on the site fairly quickly. I can, for example, post a copy of this paper as soon as I convert it to a .pdf.

As I said, the ultimate goal is to go paperless and not have to translate paper records to digital format for new projects. In keeping with this idea, I'm actually reading this paper from a digital reader, and did not waste the paper and ink to print it out. It's all very science-fictional to someone who was trained in the '70s.

I am very aware, though, of reservations about digital data (McCartney 2002:93), mainly that data repositories are threatened by (1) short term factors, like power failure, and (2) long-term factors such as "format obsolescence" and "media decay." I've had a reminder of these problems just this week. I am using the University of North Carolina's Excavating Occaneechi (Davis et al. 1997) program in my classes now, assigning students to use the electronic dig feature. Originally this program was published on CD-ROM, a real innovation. Students with Vista, however, had compatibility troubles with the CD. So this semester I assigned the web version. Well, the electronic dig link has been down for a week. UNC Press says, "We're looking into it. Thanks for notifying us of the problem."

So I am dependent on MSU's Department of Information Systems to maintain the data sharing page. They do regularly back up all data and upgrade software as a necessary part of maintaining the University's operations. On the other hand, archivists recognize that digital data are excellent for accessibility, but fragile. All digitized data will have to be backed up by hard-copy printout on acid-free paper and stored in dedicated file cabinets in the lab. The MSU Libraries are currently developing plans to incorporate an Institutional Repository, and the hard-copy files will be transferred to the Library as soon as feasible—which may mean, not during my career.

We've had to start back at the beginning, to recatalogue everything in the collection beginning with 1978, to bring it into a comprehensive, standardized and digitizable format. We're up to 1983, I think. And we've hardly begun to digitize the records. But in the end, this project will begin to fulfill the Archaeology Lab's role in stewardship of heritage resources, by preserving and presenting archaeological data to both colleagues and the public.

Mark Aldenderfer (2002) wrote a few years ago that, to take advantage of the power of computers, archaeology needs innovation in three areas: data acquisition, education (meaning training of students), and perception of the need for supporting digital presentation and preservation. Our digitization project is trying to address all three of those needs, by refining the design of digital formats for capturing archived and in-field data, by training students and incorporating the results into further training within the MSU Department of Geosciences' growing graduate emphasis in Archaeological Information Systems, and by demonstrating what resources are needed to conduct a full transfer of legacy data sets into an accessible digital form. In some ways it's the latter problem that is the most daunting: the cost of the hardware, software, and person-hours involved in designing systems, digitizing the archived data, and maintaining all of the systems for the long-term only adds to our ongoing cost of curation and overhead.

Finally, I would like to propose that we all begin a discussion of how we can make funding agencies and home institutions aware of these increasing costs of doing business. If we are required to make our data accessible in the public interest, then it is also in the public interest to help us get it done.

Thank you.

Acknowledgments

David McKenna, MSU Department of Information Services, originally set up the datasharing page that has become the MSU Archaeology Lab home page. I appreciate his help and patience.

References

Aldenderfer, Mark

2002 The Larger Context of Digital Data Dissemination and preservation in Archaeology. In Delivering Archaeological Information Electronically, edited by Mary S. Carroll, pp. 101-112. Society for American Archaeology, Washington D.C.

DAACS

2004 The Digital Archaeological Archive of Comparative Slavery (<http://www.daacs.org/>). Thomas Jefferson Foundation, Charlottesville VA. Accessed 3/5/2009.

Davis, R. P. Stephen Jr., Patrick C. Livingood, H. Trawick Ward, and Vincas P. Steponaitis
1998 Excavating Occaneechi Town: Archaeology of an Eighteenth-Century Indian Village in North Carolina. CD-ROM. University of North Carolina Press, Chapel Hill. Web version: <http://www.ibiblio.org/dig/>. Accessed 3/5/2009.

Harris, Edward C.

1989 Principles of Archaeological Stratigraphy. Second ed. Academic Press, New York.

Kintigh, Keith W.

2005 The Promise and Challenge of Archaeological Data Integration. *American Antiquity* 71(3):567-578.

Kansa, Eric C.

2007 Publishing Primary Data on the World Wide Web:Opencontext.org and an Open Future for the Past. Technical Briefs In Historical Archaeology, 2: 1-11.

http://www.sha.org/publications/technical_briefs/volume02/article01.htm. Accessed 3/5/2009.

McCartney, Peter

2002 Long-Term management and Accessibility of Archaeological Research Data. In Mary S. Carroll, ed., Delivering Archaeological Information Electronically., pp. 91-100. Society for American Archaeology, Washington D.C.

U.S. Congress

2006 *Federal Research Public Access Act of 2006*, S 2695, 109th Congress, 2nd. Session, Library of Congress, Washington, DC. <<http://thomas.loc.gov/cgi-bin/query/z?c109:S.2695>> 10 Dec 2006.