2 November 2012

Dear Members of the National Association of State Archaeologists:

We are writing to follow up on our earlier post of 24 August 2012 (see link below) to make a request for assistance, as well as provide new information about what we believe will prove to be an exciting and valuable new effort in American archaeology. The National Science Foundation has recently awarded funding for the Digital Index of North American Archaeology (DINAA) project to develop and test methods to integrate archaeological site file data from across the eastern United States into a unified information system for research and outreach (here and here). You can view the full technical proposal on the project website, and we are now assembling the team and beginning the first phase of the research design.

This project is lead from the University of Tennessee and Indiana University South Bend, along with the full partnership of the Alexandria Archive Institute's Open Context Project (one of two organizations, along with Digital Antiquity, noted as meeting the 2011 NSF standards for archaeological data management). Our research group is prepared to begin work integrating data from up to 25 states, with the goal of implementing procedures to allow the eventual integration of data from across the United States, and beyond.

Below there are several hyperlinks leading to resources and information relating to the national digital index project. Please explore them and ask any questions you wish, and share them with colleagues. If you wish to contact any of the project PIs with questions, we will gladly make ourselves available.

- Home Page and News
- Project Overview
- NSF Award announcements
- General Statement on project to NASA
- DINAA participants Data Transfer Protocols
- Link to the Technical Proposal
- Statement on security measures

This project will not distribute detailed site data since detailed descriptions are properly and legally under the control of individual states and federal agencies. Our project is directed to developing translation routines so portions of state site file databases can be integrated into larger regional and national research and management efforts. What we will need is information on how site file records are maintained in individual states, and a voluntary, one-time use of site-file data for testing and demonstration purposes. The routines and procedures we create will be made publicly available for future use. We expect this will drive a professional interest in the reuse of

site-file data and make more visible the valuable efforts of state site-file managers throughout our nation.

Here is a list of what we are interested in obtaining from your office to get started on developing the protocols to integrate your site file data with that in other states.

1. An electronic file template for the site-level data into your state's archaeological database, or (if such template does not exist) a digital or hard copy of the appropriate paper form. Scanned copies of completed forms provide an excellent example.

2. Any available information (guidelines, statutes, sites providing online availability, lists of acceptable file formats, etc.) pertaining to the completion of such a database entry as #1 (above).

3. Any list of preferred or common keywords and/or data types pertaining to the completion of such a database entry as #1 (above). Or any lists of codes, keys and dropdown menus used to fill in such a database or paper form.

4. Any guidelines specifically on the use of data for geographic information systems (GIS) including preferred coordinate systems, preferred projections, and availability of site data as point, polygon, or raster data.

5. Any information (guidelines, standards, etc.) already used by your organization specifically to create and maintain site-level archaeological data that can be integrated with data from neighboring or regional states with similar culture histories.

Please see the DINAA Participants Data Transfer Protocols for instructions on how to provide this information.

We are interested in obtaining this data from as many states as possible from east of or adjoining and immediately west of the Mississippian River, to provide coverage for as much of Eastern North America as possible. We have begun the process of contacting a number of states directly, specifically those who we contacted earlier about this projct.

Please note we are not asking for site file data at this time. We will, however, want to make use of some primary data for specific research or management demonstration projects down the line. As we have made clear, this primary site file information will be maintained securely and will not be distributed.

Once the integration routines are in place, individual states and other agencies can choose to participate, or not, in our demonstration efforts. We have commitments from people in about a dozen states to provide data for our demonstration projects so far, and so will be able to meet our project objectives with data from a large, continuous portion of Eastern North America. If possible our translation routines will include as many eastern states as possible, but our first priority will be for those states where we can make use of data to generate demonstration maps on such things as 'all NRHP eligible sites' or 'all Paleoindian' sites, and so on.

Our project will utilize security measures at the technical, procedural, and personnel levels. We understand that the security of archaeological site information must be protected for ethical as well as legal reasons. In the United States, the locations of archaeological sites represent highly sensitive data and their release could have grave repercussions. For these reasons, managing (i.e.,

permitting access to) sensitive site location data lies beyond the scope of this project, and no such information will be posted, released, or even stored on our Web-servers. Initially the sensitive data we access will only be kept on PGP-encrypted, standalone hard drives, and transferred physically between use locations for later decryption. Three of our five project principals involved with data collection and selection are RPA (Register of Professional Archaeologists) certified, with strong backgrounds in compliance and site protection. Personnel who handle sensitive data will be RPA certified or directly monitored by an RPA.

To eliminate the risk of accidental or malicious disclosure of sensitive data, our Web servers will only manage and store site location data at a very reduced level of geographic precision. Therefore, even if there is an accidental or malicious security breach, no sensitive data will exist on our servers to leak out. The exact spatial resolution we will use for public data will be negotiated with SHPO and agency personnel; this resolution is expected to be at the county scale or at ca. 20 km resolution, which have been previously accepted in earlier efforts (e.g., Anderson and Horak 1995; NADB Maps 1993). Similarly, example data products made publicly available from the project will be expunged of all sensitive data and spatial locations abstracted to at least a 20 km level of granularity.

Reduced spatial precision (and other security measures) will help protect sensitive data. Nevertheless, the public data will still permit important research programs that examine regional and large-scale geographic patterning in archaeological data. The public data will also enable powerful cross-state search and query services of great use to students, researchers and heritage managers. It will also still permit innovative Linked Open Data applications. Linked Open Data represents current best-practice in data sharing on the Web. With linked data, Web URI/URLs identify key concepts and data-points. This project will mint Web URL/URIs for all sites, making it easier to associate additional public information including objects in museum collections, educational materials, and associated research data. The project will also associate appropriate SHPO contact information with each data record to enable qualified researchers to directly obtain higher resolution spatial data from state officials. PIDBA (Paleoindian Database of the Americas) has successfully implemented similar security strategies for 22 years (Anderson et al. 2010).

We see this effort as complementing and enhancing the significance and roles of SHPOs and other agencies without endangering archaeological site security. These data will provide more exposure and better access to less sensitive data, enabling researchers, officials, the consulting community, and the public to better engage with SHPOs. Currently, few people understand the tremendous efforts SHPOs play in safeguarding the nation's heritage, and this extra public exposure can help the public better understand the richness of the heritage that lies all around them, again, without releasing sensitive information. We will also have the data under strict version control, so that it can accommodate future updates from participating SHPOs while still enabling citation and retrieval of earlier states of the record.

We are very excited to include NASA members in all general correspondence on our project as it proceeds. We welcome comment and assistance in this project from the NASA community, and look forward to hearing from and working with many of you in the months and years to come. On behalf of my co-PIs Eric Kansa (Alexandria Archive), Sarah Whitcher Kansa (Alexandria

Archive), Josh Wells (Indiana University South Bend), and Stephen Yerka (University of Tennessee), I look forward to working with you all.

Sincerely,

David

David G. Anderson, Ph.D., RPA
Professor and Associate Head
Department of Anthropology
The University of Tennessee
250 South Stadium Hall
Knoxville, Tennessee 37996-0720
865-974-2960 (o)
dander19@utk.edu
http://web.utk.edu/~anthrop/faculty/anderson.html
http://pidba.tennessee.edu/
http://bellsbend.pidba.org/

Joshua J. Wells, Ph.D., RPA
Indiana University South Bend
Department of Sociology and Anthropology
Wiekamp Hall 2288
1700 Mishawaka Ave.
South Bend, IN 46634
574-520-4434 (o)
jowells@iusb.edu
https://www.iusb.edu/sociology-anth/faculty/personnel/wells.php

Eric Kansa, Ph.D.
Lead Developer, Open Context
Researcher, UC Berkeley, School of Information
125 El Verano Way
San Francisco, CA 94127
415-425-7380 (o)
ekansa@ischool.berkeley.edu
http://opencontext.org

Sarah Whitcher Kansa, Ph.D.
Executive Director, The Alexandria Archive Institute
Editor, Open Context
125 El Verano Way

San Francisco, CA 94127
415-425-7381 (o)
skansa@alexandriaarchive.org
http://alexandriaarchive.org
http://opencontext.org

Stephen J. Yerka, MA, RPA
IT Specialist
Archaeological Research Laboratory
University of Tennessee
5723 Middlebrook Pike
Knoxville, 37921
865-974-4618 (o)
syerka@utk.edu
http://archaeology.as.utk.edu
http://pidba.tennessee.edu
http://web.utk.edu/~fac/

References Cited

Anderson, David G., D. Shane Miller, Stephen J. Yerka, J. Christopher Gillam, Erik N. Johanson, Derek T. Anderson, Albert C. Goodyear, and Ashley M. Smallwood.
2010 PIDBA (Paleoindian Database of the Americas) 2010: Current Status and Findings. *Archaeology of Eastern North America* 38:63-90.

Anderson, David G., and Virginia Horak, editors
1995 *Archaeological Site File Management: A Southeastern Perspective*. Interagency Archeological Services Division, National Park Service, Southeast Regional Office, Atlanta, Georgia.

NADB-Maps, National Archaeological Database
1993 Archeological Site Densities. State Historic Preservation Officers, National Park Service. http://cast.uark.edu/other/nps/maplib/USsitdens.1993.html