

To: Society for American Archaeology Board

From: Society for American Archaeology Task Force on Data Access and Archiving
Michael Ashley and Danielle Benden (co-Chairs), Sheila Goff, Jamie Kelly, Rachel Opitz, and Peter Nelson

Re: Report detailing Task Force Findings

The Task Force on Digital Access and Archives (TDFAA) was formed in response to an Office of Science and Preservation Technology (OSTP) memo¹ issued in February 2013 that calls for the direct results of federally funded scientific research to be made available and “useful for the public, industry, and the scientific community” to the “greatest extent and with the fewest constraints possible.” The TDFAA was asked to also consider the National Science Foundation’s (NSF) plan entitled “Today’s Data, Tomorrow’s Discoveries: Increasing Access to the Results of Research Funded by the National Science Foundation”² prepared in direct response to the OSTP memo.

Specifically, the TDFAA is charged with “examining and identifying areas of agreement, disagreement [with the OSTP memo and NSF Public Access Plan], and ways to promote best practices in data creation, sharing, curation, and (re)use in archaeology”. After a thorough review of the OSTP memo and NSF’s Public Access Plan, the six-member committee³ met telephonically on two occasions to discuss the committee charge. Committee comments were solicited and compiled by the co-chairs. The results are described herein and summarized as Action Items and Recommendations.

The OSTP memorandum has received nationwide attention and discussion as it impacts federal agencies that fund extramural research for many scientific disciplines. While the unfunded mandate is directed at federal agencies with budgets of over \$100 million annually, we praise SAA for considering the intentions of the memorandum and the broader effects on 21st century archaeological research.

The TDFAA looked at the memo through the lens of the charge, namely to assess opportunities to “promote best practices in data creation, sharing, curation, and (re)use in archaeology,” **from the perspective of both digital and physical collections and access.** With this in mind, the memo outlines major objectives for agencies that apply to SAA members.

We’ve drawn three major conclusions in our conversations and analyses:

- This particular SAA Task Force and its charge is focused principally on academic projects and archaeology, due to the particulars of the OSTP memo. Specifically, peer-reviewed research “papers” and their supporting data are the key topics we address.
- Since the vast majority of archaeology performed and supported by federal funding is largely supported by contracts and projects through cultural resources management, the

¹ https://www.whitehouse.gov/sites/default/files/microsites/ostp/ostp_public_access_memo_2013.pdf

² <http://www.nsf.gov/pubs/2015/nsf15052/nsf15052.pdf>

³ [Task Force Member bios are available here.](#)

spirit of the charge should extend to include **all** archaeology that is of interest to SAA membership, **in another phase of work.**

- The SAA should consider productive, formative ways to build pathways to publicly accessible research and data, from data creation to reuse. This will require training opportunities, policy changes, and clarifications on what can be considered ‘publishable’ and accessible, unclassified, factual research and data. In the following, we’ve outlined ways in which these pathways can be achieved.

Finally, while we have prepared this information specifically for the SAA Board, our hope is that the findings and actions detailed below will be more broadly disseminated, both across committees and to the SAA membership, as the board deems appropriate.

Action Items and Recommendations for the SAA Board of Directors

- An SAA Task Force on Data Access and Archiving in Archaeology should persist or formally be defined for the long-term. The original recommendation came out of the “Data Access and Open Archiving” memo to the Board in March 2015, and read, “the Task Force would have diverse representation from technologically-engaged archaeologists, CRM, tribal offices, and governments, as well as libraries and publishers, among other stakeholders.” This broader definition of the task force is recommended to embody the SAA more fully.
- Any discussion of data access and archiving must consider both physical and digital collections.
- SAA members - from undergraduate students to senior professionals - need access to training on the specifics of data creation, curation, and access. Fortunately, there are many experienced individuals and organizations within SAA to draw from to help create and provide this type of training. Training can be in the form of the highly successful online seminars, hands-on workshops attached to SAA meetings and regional conferences.
- The SAA Board should seek clarification from all major sponsors of archaeological research, including NSF, IMLS, NEH and NIH, regarding how each agency will account for exceptions to disseminating confidential information. Furthermore, we suggest that SAA make a concerted lobbying effort specifically focused on curation, and urge sponsors such as NSF to fund rehabilitation and enhancement projects for archaeological legacy collections, not just new projects. Specifically, we suggest SAA submit a formal letter to NSF outlining these concerns. The TDFAA or other body could draft a letter for the Board’s review.
- A minimum standard should be developed for the kinds of data that should be archived and made accessible (e.g., inventory of site collections, field maps, field notes, etc.).
- Policies for the declassification of research results and data should be established, along with educational programs of best practices for SAA membership.
- Consultation and partnership with descendent communities on the definitions of declassified data and research is essential.

OSTP Memo Assessment

There are six major sections of the OSTP memo: (1) Policy Principles (2) Agency Public Access Plan (3) Objectives for Public Access to Scientific Publications (4) Objectives for Public Access to Scientific Data in Digital Formats (5) Implementation of Public Access Plans and (6) General Provisions. Because we were directed to focus our response principally on the OSTP memo, we have organized this report based on the main sections of that memo, calling out areas of agreement, disagreement, and suggested actions. Within each section, we focus on how the OSTP memo specifically impacts the discipline, looking through the lens of data creation, sharing, curation, and (re)use in archaeology.

1. Policy Principles

The OSTP memo outlines the Administration's commitment to ensure widespread availability of scientific research data funded by federal dollars that includes peer-reviewed and digital data.

Areas of Agreement: We value the general policy principles of the OSTP memo that outlines the commitment to widely disseminate [archaeological] data in peer-review and digital formats, as they align with several of SAA's Principles of Archaeological Ethics (especially principles 1:Stewardship, 2:Accountability, 6: Public Reporting and Publication, and 7: Records and Preservation).

Areas of Disagreement: The TFDAA identified three major areas of concern with this section of the OSTP memo.

- Due to the sensitive nature of specific data in our discipline (e.g., site location details or confidential information collected for or from descendant communities), we seek clarification on how these kinds of scenarios will play out. For example, will NSF deny funding to a federally recognized tribe undertaking a research project if they choose to keep specific information confidential or will all data be subjected to dissemination?
- We are concerned that the OSTP memo completely disregards access to and use of physical collections. The focus is solely on digital curation and preservation and is lacking any discussion of the physical data including the artifacts and hard-copy associated records; Digital representations of physical objects that are created as virtual surrogates, and not as data elements, should also be addressed through the policy.
- The discipline of archaeology seems a special case to integrate with the OSTP memo for the following reasons: (1) so much of our data is tied to large physical collections; (2) we destroy sites through the process of excavation; (3) our cultural protocols are sensitive; and (4) the field is generally low-budget; when held up to other scientific disciplines, many agencies, organizations, institutions, and archaeologists will not have the funding to carry out the mandate as written.

Committee Commentary:

- We urge the SAA Board to seek clarification with all major sponsors of archaeological research, including NSF, IMLS, NEH and NIH, regarding how each of these granting organizations will account for exceptions to disseminating confidential information, and

request that each agency clearly communicate if and when omission of certain data is appropriate.

- We are optimistic that the SAA Board recognizes the importance of sponsors fully funding curation budgets through the granting process - including the physical and digital collections- and we recommend that SAA communicate this need to NSF. Furthermore, we suggest the SAA make a concerted lobbying effort specifically focused on curation, and urge sponsors such as NSF to fund rehabilitation and enhancement projects for archaeological legacy collections, not just new projects.
- While the OSTP memo is an impetus for urging federal agencies to consider greater access to data, we suggest that SAA go beyond just thinking in terms of federally-funded research and consider best practices for increasing access to any data that is generated by archaeological research that forms a part of the archaeological record.

2. Agency Public Access Plan

The OSTP memo directs federal agencies “with over \$100 million in annual conduct of research and development expenditures to develop a plan to support increased public access to the results of research funded by the Federal Government.”

Thus, the concern for SAA members is how the funding from agencies, such as NSF, IMLS, NEH and NIH, is impacted by these policies. Several mandates in this section will redefine archaeological research for the 21st Century that include developing:

- a strategy for leveraging existing archives to have a positive impact for knowledge of and access to legacy physical collections and create opportunities for digitization efforts;
- strategies for improving the public’s ability to locate and access digital data. There are relatively few digital venues for the public to readily explore archaeological projects;
- following best practices in data format and interoperability need to be encouraged to assure search capabilities, dissemination, archival storage, and long-term stewardship

Areas of Agreement: The TDFAA is in full agreement that access to non-sensitive archaeological data is an important tenet of our ethical responsibilities as archaeological professionals.

Areas of Disagreement:

- The TDFAA is concerned about the OSTP memo’s specific call for public access using existing budgetary resources. The financial burden this places on agencies, institutions, professional organizations like SAA, and independent archaeologists is cumbersome and unrealistic. Most will view this as an unfunded mandate.
- We also draw specific attention to section 2h that calls for “identification of any special circumstances that prevent the agency from meeting any of the objectives set out in this memorandum, in whole or in part.” Our previously mentioned concerns of culturally sensitive or confidential data could fall under this provision.

Committee Commentary:

- We recommend that SAA identify ways in which the financial burden of disseminating data be offset along with the ongoing costs of public access to archaeological reports and data in order to facilitate financial sustainability.
- We recommend that the SAA Board consider communicating the importance of exempting confidential data from becoming publicly accessible, especially as it may impact the ability to obtain grants and contracts if this type of data is omitted from the public record.

3. Objectives for Public Access to Scientific Publications

Section 3 of the OSTP memo calls for “results of unclassified research that are published in peer-reviewed publications directly arising from federal funding [to] be stored for long-term preservation and [made] publicly accessible to search, retrieve, and analyze in ways that maximize the impact and accountability of the federal research investment.” This section further states that:

- “Each agency plan shall ensure that the public can read, download, and analyze in digital form final peer-reviewed manuscripts or final published documents within a timeframe that is appropriate for each type of research conducted or sponsored by the agency.”
- “Shall use a twelve-month post-publication embargo period as a guideline for making research papers publicly available”
- Repositories could be maintained by the federal agency funding the research, through an arrangement with other federal agencies, or through other parties working in partnership with the agency including, but not limited to, scholarly and professional associations, publishers and libraries.
- Provides access for persons with disabilities consistent with Section 508 of the Rehabilitation Act of 1973.

Areas of Agreement:

- We agree that public access to peer-reviewed results from federally funded research (and all research for that matter) is important. The memo suggests that the repository could be maintained by the funding body or through a public/private partnership. The key is to offer the public direct access to digitally preserved research results, specifically peer-reviewed, published works.
- We are pleased to see attention paid to Section 508, which opens up the dialog for what ‘access’ means to research.
- We acknowledge that professional organizations, such as the SAA, publish for financial sustainability. The OSTP suggest a twelve-month post-publication embargo as one possible way to both offer early access to research findings for members while also allowing for public access in a relatively timely manner.

Areas of Disagreement:

- We are concerned with the portion of section 3 that reads: [that research results be] “publicly accessible to search, retrieve, and analyze in ways that maximize the impact and accountability of the federal research investment.” This may place an undue burden on SAA and other organizations that maintain publications, if responsibility for optimizing searches and maintaining content falls to the publisher.
- Section 3f(i) states that each agency shall “provide for long-term preservation and access to content without charge,” and we are concerned with how this will be attained with limited budgets.
- Section 3f(ii) states that each agency shall “[use] standards, widely available and, to the extent possible, nonproprietary archival formats for text and associated content (e.g., images, video, supporting data).” We seek clarification on what specifically each sponsor or agency’s plans will include with regard to data migration and suggested archival formats, especially since there isn’t current consensus on file formats and their preservation for the long-term.

Committee Commentary:

- We urge SAA to reach out to sponsors (e.g., NSF) to request clarification on the types of acceptable digital formats to ensure access to data. Our concern is that there isn’t a complete consensus on the current archival formats for long-term preservation, and we suggest developing a guide for specific applicable format types and best practices for migrating extant formats.
- A group of SAA members⁴ are working to build “guidelines on professional evaluation of digital scholarship in Archaeology” that articulate with the charge of this task force. We recommend they be consulted and offered an opportunity to comment further on the TFDA analysis.

4. Objectives for Public Access to Scientific Data in Digital Formats

The memo is clear that this policy applies to federal agencies with annual budgets of \$100 million dollars, but also stipulates that “extramural researchers receiving Federal grants and contracts for scientific research and intramural researchers develop data management plans”, which we translate to mean all federally funded archaeological research. Therefore, this section of the OSTP memo applies to SAA directly:

Digitally formatted scientific data resulting from **unclassified** research supported wholly or in part by federal funding should be stored and publicly accessible to search, retrieve, and analyze. Data is defined as the digital recorded **factual** material commonly accepted in the scientific community as necessary to validate research findings including data sets used to support scholarly publications, but does **not** include

⁴ The members are: Elaine Sullivan, Sarah Kansa, Joshua Wells, Heather Richards Rissetto, and Rachel Opitz.

laboratory notebooks, preliminary analyses, drafts of scientific papers, plans for future research, peer review reports, communications with colleagues, or physical objects, such as laboratory specimens.

Areas of Disagreement:

- The biggest concern we have with this section of the memo is that physical collections are specifically excluded from the discussion of preserving data, which is completely incongruent with the kinds of collections archaeologists generate.
- The specifics of what constitutes unclassified, publicly accessible data are crucial for SAA policy and action plans.
- The memo requires data management plans be developed by extramural researchers, but also allows for the inclusion of appropriate costs in proposals. A key concern of the TFDAA is how to fund and protect sensitive, classified data and documentation that is generated in virtually all archaeological projects.

Committee Commentary:

- We would like to see a strategy developed to encourage sponsors of archaeological research and federal agencies that must abide by the OSTP memo to also put policies in place for the *physical collections* and not just the digital data. When reading the NSF Plan, it appears that the physical collections may be included. At the very least, we urge SAA to seek clarification on this point, as the physical collections are a critical part of this discussion.
- We recommend that a minimum standard be developed for the kinds of data that should be archived and accessible (e.g., inventory of site collections, field maps, field notes, etc.)
- We suggest that SAA sponsor training, including an online seminar series, and educational sessions that instruct archaeologists on best practices in preparing and curating digital data.
- A clear distinction should be made between unclassified, public data - defined in the memo as factual data supporting research conclusions - and data that should not be included in order to protect people, sites, or cultural protocols. The SAA Code of Ethics serves as guide, but we recommend review of the findings of the TFDAA by the SAA Ethics Committee to assure any public data policies strive to protect these sensitivities.
- The TFDAA would like to see concise guidelines developed for SAA where cooperation with the 567 federally recognized tribes is appropriate or required. This intention should extend to indigenous communities worldwide that are impacted by archaeological research.

5. Implementation of Public Access Plans

Section 5 of the memo states that agencies with plans in place and those who do not are required to define their policies to fully meet the requirements set forth and to publish the plans publicly on the Open Government site. It is important for SAA to be aware that the OSTP/OMB will revise policies twice yearly.

Committee Commentary:

- SAA should diligently monitor key changes to agency policies that may impact SAA responses and activities.
- A key provision to the implementation of the Public Access Plan is the exclusion of digital data generated prior to the execution date (we presume the execution date to be January 2016). While this relaxes the intention of facilitating public access to federally funded research and data, **we recommend that SAA endorse a plan that encourages accessibility to legacy archaeological research and data**, pursuant to the definitions discussed in previous sections.

6. General Provisions

The OSTP memorandum defines the unfunded mandate for agencies with budgets of at least \$100 million / year. The General Provisions section clarifies that the memorandum does not create benefits nor intends to undermine the authority of its intended audience.

Committee Commentary:

We are concerned about the logistics of how each agency that must comply with this memo will move from plan to implementation without additional funding.

Important Updates to the OSTP Memo and NSF Requirements That Impact Archaeology

Updates to the OSTP memo are available here:

<https://www.whitehouse.gov/administration/eop/ostp/library/publicaccesspolicy>. The latest update was July 29, 2015. In the April 8, 2015 update, the 13 agencies with public plans were identified, including NSF. The NSF Plan⁵ updated in June 2015⁶ stipulates that all funded projects effective January 2016 meet the following requirements:

Either the version of record or the final accepted manuscript in peer-reviewed scholarly journals and papers in juried conference proceedings or transactions **must**:

- Be deposited in a public access compliant repository designated by NSF;
- Be available for download, reading and analysis free of charge no later than 12 months after initial publication;

⁵ https://www.nsf.gov/news/special_reports/public_access/

⁶ https://www.nsf.gov/publications/pub_summ.jsp?ods_key=opengov

- Possess a minimum set of machine-readable metadata elements in a metadata record to be made available free of charge upon initial publication;
- Be managed to ensure long-term preservation; and
- Be reported in annual and final reports during the period of the award with a persistent identifier that provides links to the full text of the publication as well as other metadata elements.

Areas of Agreement:

- Providing sustainable, public access to federally funded research with machine-readable metadata is a reasonable and worthy goal.

Areas of Disagreement:

- The Archaeology Division of NSF⁷ “does not sponsor, endorse or have an official arrangement with any data archive.” The NSF should provide guidelines for open access, not prescriptive designations for *official* repositories.
- SAA may wish to provide guidance to the embargo period in their Data Management Plan Guidelines⁸ or to discuss options with SAA membership.
- There needs to be clear guidance on what is considered minimally acceptable, machine-readable metadata elements, that are especially useful in an archaeological context.
- It is often difficult to provide permanent IDS and digital object identifiers (DOIs) on works in progress. We would suggest this be part of the final phase of grant work.

⁷ https://www.nsf.gov/funding/pgm_summ.jsp?pims_id=11690

⁸ https://www.nsf.gov/sbe/bcs/arch/SAA_Data_Management_Plan_Guidance.pdf