Department of Defense

Legacy Resource Management Program

PROJECT NUMBER 13-711

Evaluating a Cooperative Approach to the Management of Digital Archaeological Data

Sara Rivers Cofield and Jodi Reeves Flores

December 2014

Evaluating a Cooperative Approach to the Management of Digital Archaeological Records

Department of Defense Legacy Resource Management Program
Project No. 13-711

Sara Rivers Cofield
Curator of Federal Collections
Maryland Archaeological Conservation Laboratory, Jefferson Patterson Park and Museum

Jodi Reeves Flores
Digital Curator, Center for Digital Antiquity
Council on Library and Information Resources/ Digital Library Federation Fellow

Additional Contributors

Francis P. McManamon
Executive Director, Digital Antiquity

Adam Brin
Director of Technology, Digital Antiquity

Grant Snitker
Digital Curator, Digital Antiquity

Chelsea Walter
Digital Curator, Digital Antiquity

Michael A. Smolek
Cultural Resource Manager
Naval Air Station Patuxent River

Amanda Vtipil
Curator of Education
U.S. Army's Women's Museum

December 31, 2014

Abstract

The Department of Defense (DoD) needs efficient access to data from past archaeological investigations at its installations in order to avoid sudden, unpredicted site discoveries that delay mission-oriented activities, programs, and projects. The ECAMDAR project is a test case designed to evaluate whether and how an online repository for digital archaeological and cultural resource management (CRM) data and information developed and managed by the Center for Digital Antiquity (Digital Antiquity) can fulfill this need. Digital Antiquity's online repository is tDAR (the Digital Archaeological Record).

With funding from Defense Legacy, digital files from archaeological projects at 23 DoD installations in Maryland and Virginia have been uploaded to tDAR, which is accessible through the tDAR website via the internet. This study finds that through tDAR, Digital Antiquity is able to ensure long-term preservation and accessibility of digital archaeological records while maintaining security by enabling DoD CRM officials to control and limit access to sensitive files. As files were uploaded for this project, DA-tDAR worked with physical archaeological collections repositories to develop procedures that address the backlog of digital files that are held locally at the repositories, but not currently in a viable long-term digital archive. These procedures may be applied to other DoD installations with pre-existing digital records. It is also important for current and new archaeological projects that the digital files created are placed in a digital archaeological data repository so that no further backlog of digital data develops. We suggest tDAR for this purpose.

This project demonstrates that it is more cost effective for the DoD to pay one-time fees for uploading of digital files to Digital Antiquity than it would be for the DoD to establish and maintain individual specialized departments or staff for the management of digital archaeological data at the installation or regional level. Finally, this study finds that the DoD should consider partnership with Digital Antiquity for the preservation and management of digital files generated by current and future archaeological projects.

Table of Contents

List of Figures

List of Tables

1. Introduction

Sara Rivers Cofield

1.1 Project Summary

There is currently no system at the DoD Service or Command levels for preserving and disseminating digital data generated by archaeological work on military installations. Essential documentation of archaeological investigations such as surveys and excavations increasingly are created and stored in digital form only. Digital files are vulnerable to corruption, hardware failure, and format obsolescence if they are not properly maintained, preserved, and migrated. This is a problem because the DoD's considerable past and ongoing investment in managing its archaeological resources is at risk. Without proper management and long-term accessibility and preservation of digital data, the results of expensive archaeological work may be lost altogether, wasting money and leaving installations unable to factor significant archaeological resources into their activities, developments, and training plans. This project explores a possible solution to the problem.

The Digital Archaeological Record (tDAR) is an international digital repository for the records related to archaeological investigations and resources, as well as other CRM data and information. tDAR's use, development, and maintenance are governed by the Center for Digital Antiquity (Digital Antiquity), an organization dedicated to ensuring the long-term preservation of irreplaceable archaeological data, broadening access to these data and encouraging the appropriate use of these data to learn about the past and manage archaeological resources effectively. Digital Antiquity (http://www.digitalantiquity.org/) is part of Arizona State University (ASU). Digitial Antiquity is designated as a university center so that ASU provides administrative, legal, and high-level supervisory support for the Center's activities. In addition, Digital Antiquity also has a Board of Directors, independent of the university administration, which is composed of well-respected archaeologists representing a number of institutions (the University of Arkansas, ASU, the Pennsylvania State University, Statistical Research, Inc., Washington State University, the University of York Archaeology Data Service), as well as experts in digital library information management, not-for-profit organization management, intellectual property law, and finances. Digital Antiquity's business plan and operating methods are explicitly designed to ensure the long-term financial, technical, and organizational sustainability of tDAR.

The purpose of the ECAMDAR project is to evaluate Digital Antiquity as a potential partner for the management of all of the DoD's digital archaeological records and tDAR as an appropriate repository to meet the needs of the DoD. While the study does not look at digital files from installations nationwide, it does adopt a regional approach through the inclusion of data from two archaeological repositories: the Maryland Archaeological Conservation

Laboratory (MAC Lab) at Jefferson Patterson Park and Museum (JPPM), and the Regional Archaeological Curation Facility (RACF) at Fort Lee. Together, the MAC Lab and the RACF curate collections from 25 DoD installations, and the cultural resource officers responsible for these installations all supported the inclusion of their digital archaeological data in the project (Table 1).

Table 1: Participating installations from the Maryland Archaeological Conservation Laboratory (MAC Lab) collections and the Regional Archaeological Curation Facility (RACF) collections. Two of the 25 installations had no digital records whatsoever, so only 23 of the installations were able to contribute to the ECAMDAR project.

Repository	Installation	Data (nearest MB) Submitted	Approx. # Files Submitted	Point of Contact
MAC Lab	Naval Air Station Patuxent River	1229	2327	Michael Smolek
MAC Lab	Naval Air Station Patuxent River, Webster Field Annex	1462	1202	Michael Smolek
MAC Lab	Point Lookout	1286	224	Michael Smolek
MAC Lab	Bloodsworth Island	1	6	Michael Smolek
MAC Lab	Solomons Naval Recreation Center	276	111	Michael Smolek
MAC Lab	Naval Support Facility, Indian Head	145	425	Thomas Wright
MAC Lab	Naval Observatory	74	43	Julie Darsie
MAC Lab	Potomac Annex	10	37	Julie Darsie
MAC Lab	Washington Navy Yard	354	93	Julie Darsie
MAC Lab	Joint Base Anacostia Bolling	790	346	Julie Darsie
MAC Lab	Nebraska Avenue Complex	126	61	Julie Darsie
MAC Lab	Walter Reed National Military Medical Center	121	67	Julie Darsie
MAC Lab	U.S. Naval Academy	394	217	Kimberly Hickey
MAC Lab	USNA Dairy Farm	177	64	Kimberly Hickey
MAC Lab	North Severn	386	124	Kimberly Hickey
MAC Lab	U.S. Army Garrison Aberdeen Proving Ground	1666	1625	Mark Gallihue
MAC Lab	U.S. Army Garrison Adelphi Laboratory Center	934	638	Jim Krake
MAC Lab	Fort George G. Meade	1448	1003	Jerry Glodek
MAC Lab	Fort Detrick	20	30	Alfred Lynn Hoch
RACF	Fort Lee	1331	665	Amy Wood
RACF	Fort Monroe	1745	146	Amy Wood
RACF	Fort Eustis	N/A	N/A	N/A
RACF	Quantico Marine Corps Base	862	516	Kate Roberts
RACF	Dahlgren Naval Support Facility	N/A	N/A	N/A
RACF	Fort A.P. Hill	7813	1653	John Mullins

Determining whether Digital Antiquity would be an appropriate partner and tDAR a good tool for DoD-wide management of archaeological digital records requires the consideration of four important questions:

1) How would partnership with Digital Antiquity using tDAR support the DoD's military mission?
2) What are the potential security risks in using tDAR, and can these risks be mitigated?

3) Is the use of tDAR cost effective?
4) Why should the DoD create a partnership with Digital Antiquity and use tDAR instead of its own internal information technology (IT) resources and staff?

The following sections of this report summarize the project background, methods, and results, including the results of a survey conducted to solicit comments from participating installation points of contact. Case studies drawn from the experiences of archaeological curators, Digital Antiquity staff, and installation points of contact will appear throughout the report to illustrate key points.

The results of the project suggest that using tDAR as a digital archaeological archive is a cost effective and secure method of preserving DoD digital archeological records. Furthermore, tDAR offers a level of access to archaeological records for installation CRM staff that traditional curation cannot provide, and this ease of access is generally not allowed or desirable for internal DoD computer systems. Fast, secure access to archaeological records facilitates efficient planning and impact-assessment for mission-oriented activities, projects, and construction that could impact archaeological resources.

Because the current project argues that Digital Antiquity is an organization that the DOD as a whole should consider using for archaeological data management and preservation, this report draws from its findings to develop language and templates that may facilitate the adoption of Digital Antiquity's services, such as:

1) **Justification for services:** Explanation of the services and the laws requiring such services. This justification may be needed to facilitate the procurement process.
2) **Language for scopes of work:** Standardized language can be included in scopes of work for archaeological projects to ensure that the digital records generated are addressed.
3) **Sample fees:** Digital Antiquity has a fee schedule in place and can offer quotes for addressing existing digital data, uploading files from new projects, and a variety of other digital curation services.
4) **Logistics:** Procurement procedures vary by installation, so logistical possibilities are described for the inclusion of digital archiving in CRM SOWs, curatorial repository SOWs, or directly between client installations and Digital Antiquity.

The suggested language developed as part of the ECAMDAR project is included in appendices where it can be copied for use in developing new policies that will ensure the protection and availability of archaeological data generated by the DoD.

1.2 Authorship

The ECAMDAR project has been a team effort, but since the team includes both repositories and installation PoCs who tested tDAR, as well as the tDAR staff being tested, it is necessary to

clarify authorship throughout this report to make it clear that the findings are not simple self-promotion on tDAR's part.

The primary leads initiating the ECAMDAR project were Sara Rivers Cofield, Curator of Federal Collections at the MAC Lab, and Amanda Vtipil, a Versar, Inc. employee contracted with Ft. Lee to act as the Curator for the RACF. Not long after the ECAMDAR project was funded, Vtipil changed positions, so her contributions were primarily limited to the project preparation. In her stead, Amy Wood became the contact for the RACF as the Ft. Lee Cultural Resource Manager. Wood also changed positions in October 2014, however, so Rivers Cofield acted as the author of the report on behalf of both of the curation repositories.

Points of contact from the participating installations contributed to the ECAMDAR project by working with Digital Antiquity and their curatorial repositories to provide background information and feedback through surveys, e-mails, and phone conversations. Their point of view is therefore incorporated throughout the report.

Finally, Digital Curator Jodi Reeves Flores acted as the lead for the ECAMDAR project on behalf of Digital Antiquity, and is a lead author on this report along with Rivers Cofield. By necessity, the portions of this report relating to Digital Antiquity and tDAR as an organization, the technical foundation of tDAR's work, and the operational details of how tDAR took in the DoD data included in this project, were all authored by Reeves Flores with input from Digital Antiquity's Executive Director, Francis P. McManamon and tDAR's Director of Technology, Adam Brin.

The different authors are listed for each section to clarify the point of view presented (curator/installation vs. Digital Antiquity). However, as the report will make clear, even the people responsible for evaluating tDAR approached this project with optimism and the assumption that using tDAR would probably be a good idea. There was always a possibility that tDAR would not live up to expectations, so critical analysis was applied throughout the experiment, but ultimately, tDAR was adaptable enough to address any criticisms. All sections of the report therefore reflect a pro-tDAR point of view regardless of the author.

2. Background

2.1 Defining the Problem -Sara Rivers Cofield

Archaeology performed by the DoD to comply with laws and regulations results in the need to curate physical artifacts uncovered during excavation, related paper documents and records, and digital data and files that record the site(DoD 2005: Appendix A). Documentation is essential because excavation is destructive. It cannot be redone if the original descriptive documents and analytical results for an archaeological project are lost. Without both artifacts and documentation, the contextual information needed to interpret the site, undertake further study, and make collections meaningful is lost.

For example, archaeologists take photos of excavations to record soil layers and features that are destroyed by the act of excavation (Figure 1). The data captured in these images is essential for understanding a site. Recent years have seen a decline in film photography that produces archival-quality, paper-based, photo printing. Instead, archaeological photos increasingly exist in digital form only, making it imperative that careful digital data management be provided in order to prevent the loss of information about archaeological sites. This information has been collected for the public benefit at great expense to the DoD.

Similar problems exist for digital files such as artifact inventories, reports, and maps. The software used to generate these files and the hardware used to store the files becomes obsolete as rapid changes in technology take place. Professional digital archiving practices and procedures that ensure the long-term preservation of digital documents, data sets, images, etc., such as the use of appropriate standardized file formats, are needed to avoid technological stagnation and information loss. Old files must be diligently migrated and automatically and systematically monitored to detect and remove obsolescence and corruption. As the influx of digital files swells, and the files themselves age, the need for professional and dedicated digital archivists becomes imperative.

The proportion of archaeological records in digital form is already substantial and is increasing exponentially (Figure 2). Some of the records of contemporary archaeological and CRM investigations, e.g., geospatial data sets (GIS and GPS data) and artifact or landscape (LiDAR) scan files, exist only in digital formats. Curatorial studies may also result in files that are exclusively digital, such as 3D scans. For example, the Virtual Curation Laboratory at Virginia Commonwealth University is a 3D scanning project that generated enormous digital files using Defense Legacy Program funds (Haynes and Means 2011; Means 2013). Just as artifacts need

a physical repository, the 'virtual' artifacts created through this effort will need a digital repository in order to be viable long-term.

Figure 1: Excavation of archaeological sites is destructive. As part of excavation, soil layers are removed along with evidence of human habitation such as post holes, storage pits, and artifact clusters. Archaeologists therefore document each step of excavation with maps, drawings, and photos of the soil differences exposed. Soil profiles (top) show differences in color indicative of features where human hands have altered the landscape. As sites are excavated, photos are taken to show the exposed features, the extent of excavation (bottom left), and *in situ* placement of notable artifacts (bottom right). Images such as these, taken during the VXX helipad project (see Case Study 6), are all that remains of sites that are later impacted by development, and these images exist only in digital form.

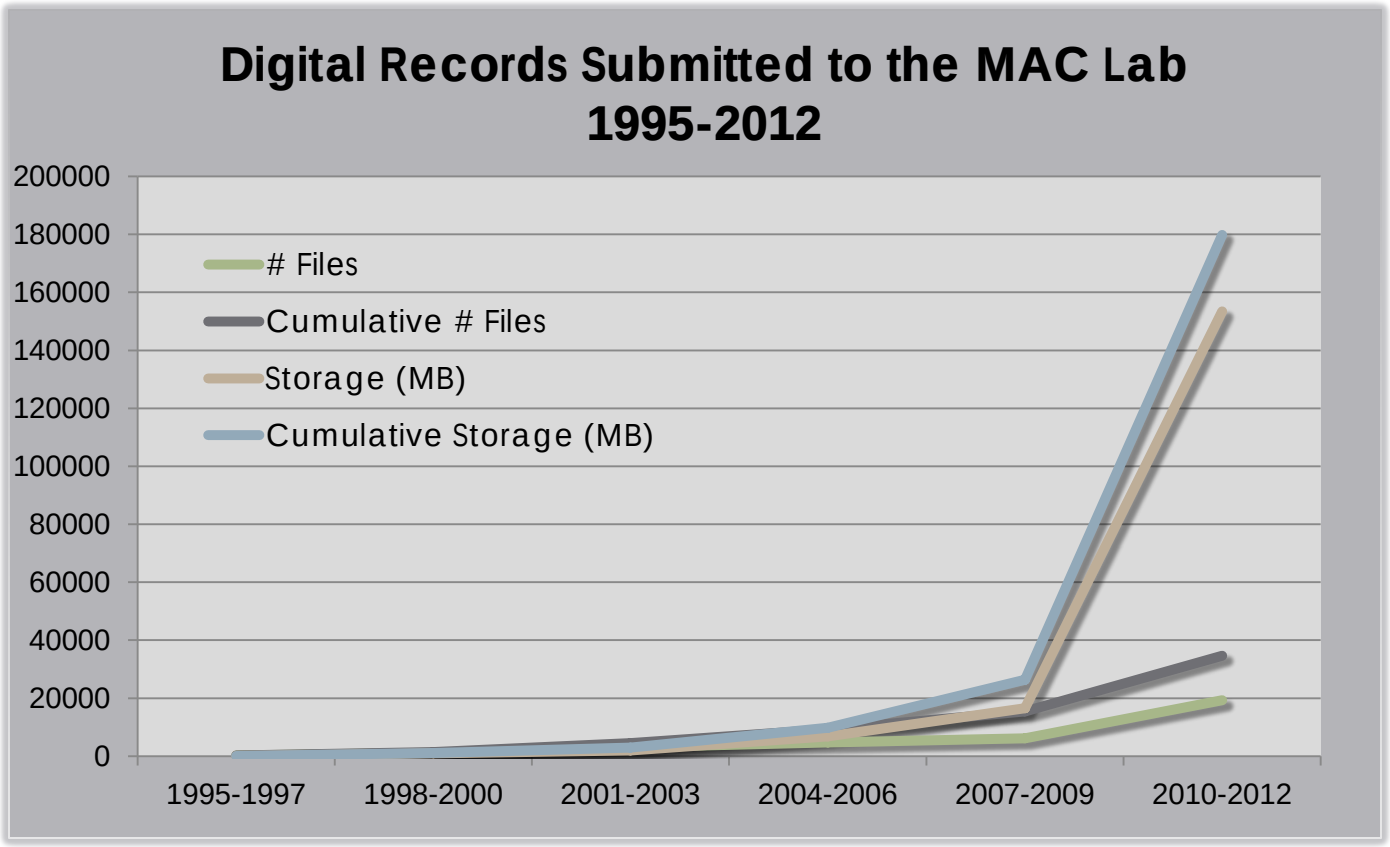

Figure 2: Total digital records submitted to the MAC Lab from 1995-2012. Recent years have seen a dramatic increase both in the number of files submitted and in the storage space needed to keep them.

Federal regulations, in particular, but not exclusively, 36CFR§79, require that archaeological collections, including artifacts and associated digital and paper records, must be properly curated and preserved to ensure that the public interest in cultural resources is protected even if sites are destroyed by DoD activities. A legal review conducted by Cultural Heritage Partners, PLLC (2012) found that the laws that mandate archaeology and the curation of resulting collections apply to digital records and mandate that these records be accessible and preserved in perpetuity (Appendix A, Part 2).

Previous Defense Legacy projects have generated recommendations for processing DoD archaeological collections, including electronic media, and the guidance offered would ensure the long-term preservation of data (Griset and Kodack 1999; Sagebiel et al. 2010). However, in the short-term it is not necessarily the case that curatorial repositories are capable of fulfilling the suggested requirements, and in the long-term the processing recommendations are unlikely to remain relevant as technology changes.

In 1999, Griset and Kodack (1999:62-67) said that electronic media should be stabilized and organized for accessibility. Emphasis was placed on ensuring that files could open without errors or viruses and that storage media (especially CDs) were properly labeled. Ultimately,

however, Griset and Kodack (1999:67) argued that digital data is not permanent, it cannot replace "original" documentation, and it should be viewed as a convenience for access, not as a long-term solution for data storage. Instead, hard copies of all digital files must be printed and retained as a more stable record. That recommendation was made before the exponential increase in the use of digital media that has occurred over the past 10 years. More recently, Sagebiel et al. (2010) expanded on Griset and Kodack's (1999) guidelines by offering more detail about how to preserve electronic media (Figure 3). Hard copies are still emphasized as a necessary accompaniment to the digital files, but there is also more information about data management.

Electronic media

- Should be checked for viruses and errors.
- Both an electronic and hard copy should be made.
- Hard copies of all files should be made on acid free paper and labeled with the number of the corresponding electronic media, file name, and the software used.
- Magnetic media such as diskettes, floppy disks, hard drives, DAT, and similar tapes are susceptible to magnetic fields, static, dust, and humidity. Optical media (CDs and the like) are more stable.
- Do not use rewritable disks.
- Electronic media should be numbered, labeled directly, and further described in the catalog or finding aid.
- Do not use stickers to label disks, use pens designed specifically for labeling disks.
- For each electronic file, record the format, version of the software used, date of creation, creator's name, file relations, and database structure and related scripts or macros (Sullivan and Childs 2003:37–38).
- The type of software used should be carefully considered for compatibility and long-term utility (Sullivan and Childs 2003:37–38).
- Electronic files will need to be transferred to new media periodically. Most have only a 10-year life span because of hard and software compatibility issues.
- Document each transfer or update (Sullivan and Childs 2003:37–38).
- Use publicly defined file formats rather than proprietary formats. For example, do not use Word's Doc format but RTF or uncompressed TIFF format (Childs and Corcoran 2000, www.nps.gov/history/archeology/collections/Field_02.htm; Eiteljorg 2004).
- Do not embed files such as photographic files within a text file because that complicates data migration (Eiteljorg 2004).
- Avoid using glossy or compressed formats for photographic files (Childs and Corcoran 2000 www.nps.gov/history/archeology/collections/Field_02.htm).
- Disks should be placed in polypropylene containers.
- Corrosion Intercept® pollutant scavenger inserts may be added to protect CDs from gases and fumes in the environment (Brady et al. 2006 http://www.sha.org/research_resources/conservation_faqs/curation.cfm).
- Should be stored in a cool, dry, environment with 20–30% humidity.
- A good resource is the Library of Congress's "Digital Preservation" webpage (http://www.digitalpreservation.gov/).

Figure 3: Extract from Sagebiel et al. (2010: 36) listing the recommended preventive conservation techniques for preserving and storing archaeological data on electronic media.

While the bullet points outlined by Sagebiel et al. (2010: 36) for preserving digital data are valid, their implementation is problematic. Some of the points summarize essential processes that are much more complex and difficult in practice than the list suggests. For example, the report says that electronic media, "should be checked for viruses and errors," and, "electronic files need to be transferred to new media periodically" (Sagebiel et al. 2010:36). Both statements are absolutely true, but ongoing assessments of file viability and periodic upgrades require specialized IT skills. Curators could open each file manually, check the CDs, migrate formats as needed, and copy to new CDs in case the old ones deteriorate, but the work would increase exponentially with each new collection acquired. Such rigorous data monitoring would require dedicated digital curators and well-supported technological infrastructure; assets that archaeological repositories are unlikely to have. As Wendy Bustard (2000:12), a Curator for Chaco Culture National Historical Park, pointed out in 2000, "Migrating data files to new media every five years or so is a worthy goal, but one that may not be realistic, given other curatorial concerns and crises."

Other parts of the guidelines are subject to rapid change. Even though Sagebiel et al.'s (2010:36) recommendations are only four years old, data storage has already changed dramatically as CDs have largely been replaced by flash drives and "cloud" storage has entered the mainstream. Formats can become obsolete within a few years, software may not prove to have long-term viability no matter how carefully chosen, and the ease with which files are migrated is subject to change as technology changes. Like many curators, Bustard expressed an interest in having national standards for migration, verification, and preservation of digital media, but at the same time she acknowledged that, "The media change so quickly that long-term preservation studies are non-existent and would be largely irrelevant" (Bustard 2000:12). In short, digital technology is a dynamic force that will not be predictable in the same way that archival bags and tags are predictable, so guidelines for digital data management will require revision much more often than other curatorial standards.

This is undoubtedly why archaeologists and collections managers rely heavily on having a hard copy of everything. Acid-free paper is a known quantity, it is stable, and it is something curators can predictably preserve. However, reliance on hard copies is problematic as well. The long-term stability of printed records depends on the quality of the printer and ink, which can vary greatly and is difficult to regulate. More importantly, however, digital files are undeniably "original" records of excavations, not just convenient formats for access. Photography is almost exclusively digital, reports and figures are computer-generated, and field forms are increasingly created using portable tablets. In other words, hard copies are not necessarily the original documents anymore, and in an increasingly paperless society, they are unlikely to make a big comeback in the near future.

As early as 1997, Federal courts ruled that e-mails preserved by the National Archives and Records Administration (NARA) in accordance with Federal law could not just be kept as hard copies. Instead the ruling stated that, "electronic communications are rarely identical to their paper counterparts; they are records unique and distinct from printed versions of the same record" (quoted in Chittenden 1998:17). The preservation of the integrity of a digital file therefore must include the maintenance of its functionality as an electronic document. This means that the long-term viability of original digital records is a concern regardless of whether or not a hard copy is made.

Unfortunately, curatorial repositories built and staffed for the preservation of physical archaeological collections are not equipped with full-time experts in IT or the professional procedures and practices needed to ensure that digital information is preserved and remains useful (Case Study #1). Digital files require different kinds of care and procedures than physical collections to ensure that they are properly preserved and accessible for appropriate uses (Table 2). The nature of digital curation is not necessarily more complicated or expensive than physical collections, but it is specialized and the DoD needs to take affirmative steps to ensure that the archaeological data about their resources and from their projects are deposited in an archive or repository where the expert care, principles, standards, and techniques of digital curation are followed.

The two repositories participating in this project, the MAC Lab and RACF, represent leading professional standards in archaeological curation. Both boast relatively new construction of facilities designed for the needs of artifacts and paper records. Neither, however, has a digital archivist on staff. At the MAC Lab, digital records are copied to a local area network by accession number and they are also stored on archival CDs. All files are backed up on tapes which have daily, weekly, and monthly back-up cycles. The data is therefore protected from immediate loss if the network crashes. However, this system does nothing to check files for corruption, nor are files migrated to new standard file types as the software used to produce them becomes obsolete. It is not uncommon to find that when old digital files are needed for research, they no longer open. The existing system does not provide an easy means of sharing the data, even with the installations that are responsible for the resources from which the data are derived.

The RACF recently amended collections standards to require depositors to submit electronic copies of all associated documents (RACF 2012). This has greatly increased the number of digital records housed at the facility. Currently all digital files are copied to a 1TB external hard drive, a policy that resulted in part from the problem of relying on the stability of CDs as a storage medium (Case Study #2). Use of the external hard drive is limited to non-networked computers though, because the use of USB connections is not authorized on computers on Ft. Lee's network. Subsequently, access to and sharing of the digital records is limited. The information contained on the external hard drive is backed up on CDs which are stored at an

off-site location. Similar to MAC Lab practices, files are not regularly migrated or checked for corruption.

CASE STUDY #1
Curation Desperation

Sara Rivers Cofield
Curator of Federal Collections
Maryland Archaeological Conservation Laboratory

Many years ago I inherited an ex-employee's drawer full of 3.5" floppy disks because they contained the working files for a number of archaeological projects on Navy property. As the Curator of Federal Collections, it was my responsibility to ensure that all documentation of the Navy projects would be preserved in-perpetuity. So before our last computer with a 3.5" floppy drive died, I uploaded the contents of all of the disks to our server.

Unfortunately, it was already impossible to open many of the files because we no longer had the software programs that generated them. Additionally, sometimes moving the files rendered them unusable. For example, mapping programs like Surfer could no longer find the component files needed to generate the original maps. No one left any notes with the disks to indicate which files went together, and the file names were so abbreviated that they were not helpful in determining what each file was for. In frustration, I realized quickly that our facility may be "state of the art" for storage and conservation of artifacts, but we are not at all equipped to deal with digital file preservation. The skill set is totally different, and we were still operating as if film photography and paper records were the norm for site documentation.

The IT staff in our parent offices initially wanted us to clean out our old files, not understanding that we were responsible for in-perpetuity curation. Even when we made that clear, it still is not their expertise, since archiving digital records is different than general maintenance of a workplace IT system.

Eventually I discovered that the library sciences are the best resource for digital archiving, and the ideal solution would be to hire a specially trained full time digital archivist. Maryland State government has been cutting positions and budgets in recent years, however, so the outlook for getting a digital archivist is bleak. Plus, the longer we wait for the staff position, the more inundated we are with a backlog of unmanaged files. In the meantime, how could I answer a client if they asked why the records they had deposited with me were no longer viable? Still, we need to fulfill our stewardship responsibilities, so I continued to look for a solution.

I had heard of tDAR and learned that it was the best option available. My options were to either ignore the problem or seek funding to test tDAR with my Federal records. Even if the project is finite and does not address our whole facility's backlog of digital files, at least we could work with professionals to develop standards that might be adopted for the future submission of digital records. That is how this Defense Legacy project developed. By giving tDAR a try, we could do something to cover this gap in our ability to care for collections.

Table 2: Examples of the requirements for the professional preservation and management of digital data versus artifacts, paper, and photo documentation.

Requirements	Digital Data Curation	Artifact/Paper Curation
Maintain a storage facility within the collection's region or State of origin for ease of access.		✓
Monitor the storage environment for relative humidity, temperature, mold, pests, and other agents of deterioration.		✓
Recognize the deterioration of materials and seek conservation as needed.		✓
Maintain secure storage and handling methods to prevent loss from theft or artifact breakage.		✓
Offer work spaces for the physical inspection of artifacts by curators and researchers.		✓
Facilitate artifact loans and exhibits as appropriate.		✓
Address requests for repatriation in compliance with NAGPRA.		✓
Conduct periodic inventories and inspections of artifact boxes and boxes of associated paper records.		✓
Facilitate public access to collections while protecting confidential information such as site locations.	✓	✓
Regularly and systematically check digital files to ensure that no deterioration has occurred. If file deterioration is detected, take steps to remedy it.	✓	
Periodically migrate and/or refresh the digital files to provide for their long-term accessibility and preservation.	✓	
Plan for obsolete technology.	✓	
Maintain files in open and preferable formats, and accommodate new industry standards for archaeological information.	✓	
Store rich, descriptive metadata with each digital object.	✓	
Ensure that all materials deposited are properly backed up.	✓	

CASE STUDY #2
CDs ≠ Stable Archives

Amanda Vtipil
Curator of Education
U.S. Army's Women's Museum
(Former RACF Curator)

In the U.S. Army, CDs are the common device for digital storage, as use of USB ports is not authorized on government computers. Even before the Fort Lee Regional Archaeological Curation Facility's Collection Standards were changed in June 2012 to require the submission of digital records on archival quality CDs, the collections held there contained a large number of CDs. There was no comprehensive inventory of the digital files though and the CDs were in various states of preservation – wrapped in notebook paper, slipped into a box of artifacts, in protective cases, etc.

For this project all boxes were systematically inventoried for digital records. All digital files stored on CDs were copied to our external hard drive and backed up again on more CDs. During this process, several CDs were found to be no longer viable. Unfortunately, the information stored within these CDs is essentially lost to time. While we tend to think of CDs as relevant and safe forms of digital storage we need to recognize the danger of relying solely on them as a means of preservation in perpetuity.

According to the U.S. National Archives and Records Administration (NARA 2014):

> CD/DVD experiential life expectancy is 2 to 5 years even though published life expectancies are often cited as 10 years, 25 years, or longer.... Life expectancies are statistically based; any specific medium may experience a critical failure before its life expectancy is reached. Additionally, the quality of your storage environment may increase or decrease the life expectancy of the media. We recommend testing your media at least every two years to assure your records are still readable (NARA 2014).

While CD/DVDs may last longer than 2-5 years, the NARA (2014) figures indicate that counting on the medium to last longer is a gamble. Add digital obsolescence on top of this and reliance on CDs as a media storage device is even riskier. That is why it is so important to consider other means of digital preservation like The Digital Archaeological Record (tDAR). As a field we need to recognize the need to create good records, deposit them with trusted repositories and ensure the future accessibility of these records.

Curatorial staff versed in the care of artifacts and paper records are well-suited to the long-term preservation of analog data in that they are trained to keep physical objects safe and to recognize deterioration as it arises (Figure 4). Digital media, however, shows no physical signs of decay and inexorably becomes unstable and unusable. It must be checked on a regular basis, replaced if it deteriorates, and migrated to improved standard file formats as these develop and improve information management. At present, there is no DoD-wide system for

long-term preservation and management of digital archaeological data. Federal agencies in different parts of the U.S. already have difficulty finding 36CFR§79-compliant repositories, and the requirement of digital data management standards for repository qualifications aggravates the problem (Bawaya 2007; Bustard 2000; Childs 1995, 2004; Kodack and Trimble 1993; Thompson 1999).

Fortunately, technology allows digital data to be stored, maintained, and accessed remotely, eliminating the need to develop, staff, and operate multiple local or regional repositories in order to ensure accessibility. The DoD could therefore consider using a centralized digital repository to meet its archaeological digital data management responsibilities.

Figure 4: Archaeological repositories are typically equipped to ensure proper packaging and care of artifacts (top left), and photos, and paper records (bottom left), but simply adopting archival storage for computer disks (above) is not enough to ensure the longevity of the data they contain.

2.2 Cooperative Curation -Sara Rivers Cofield

Historically, the DoD has chosen to partner with non-DOD institutions to curate its archaeological collections. Instead of looking at DoD facilities as possible repositories, the Army Corps of Engineers (USACE) has specifically evaluated non-military curation facilities throughout the country as potential partners in an effort to maximize the research value of

collections and minimize the DoD's need to focus on activities that are important, but not primary to the function of the military (Bustard 2000; Felix, et al. 2000; Langness, Marino, and Van Arsdale 2000; USACE 1999). According to the Mandatory Center of Expertise for Curation and Management of Archaeological Collections (MCX-CMAC):

> Military installations or other DoD/USACE facilities were not [studied as] potential partners since these institutions' primary mission is not the long-term curation of archaeological collections; their primary function is not archaeological collections management, staff are not always available to care for the collections, and public education and use of the collections cannot always be assured. [USACE 1999:ix-x]

Upon implementation, the creation of curatorial partnerships has proven to be cost effective as well as beneficial for research (Futato 1996; Hanniball 2000; Rivers Cofield 2005). By depositing collections with universities, for example, installations can eliminate significant infrastructure and staffing costs while putting collections in locations that promote their use in conjunction with academic resources such as DNA laboratories, libraries, and faculty expertise (Futato 1996). While a potential drawback of cooperative curation is loss of control on the part of the DoD, each installation can decide to stay involved as much as they see fit through a carefully negotiated Memorandum of Understanding.

The cooperative curation model has been tested and has proven to be successful and cost-effective for the DoD's archaeological collections, so it is reasonable to expect that a similar approach could work for digital archaeological records. Just as the MCX-CMAC office evaluated curatorial repositories for physical collections (artifacts, photos, paper records), this project is designed to evaluate a repository for the digital files associated with DoD collections.

Unlike the USACE curation options projects, the ECAMDAR project is not studying a variety of repositories nationwide. Instead, only one repository— tDAR— is being evaluated. The reason for the limited scope of the ECAMDAR project is simple; tDAR is the only digital archaeological repository in the U.S. at this time. Ideally, it would be beneficial to evaluate a number of options and make recommendations, but it is too risky to let existing digital data go neglected while waiting around for more choices to present themselves, especially when a repository exists already that could work for DoD digital archaeological data nationwide.

2.3 tDAR: An Option Worth Testing –Francis P. McManamon

The Digital Archaeological Record (tDAR) is an international digital archive and repository that houses data about archaeological investigations, research, resources, and scholarship. tDAR provides researchers new avenues to discover and integrate information relevant to topics they are studying. Users can search tDAR for information about digital documents, data sets, images, and other data resources. For a large percentage of these digital objects, registered users can download a copy of the digital document or other file, unless the digital object has

been marked "confidential." The choice of whether or not to designate a file as "confidential" in tDAR is made by the individual who uploads the file or organization that authorizes the deposit of the file in tDAR.

The repository encompasses digital data, documents, and images derived from ongoing archaeological research, as well as legacy data derived from more than a century of archaeological research. Since September 2010, tDAR has been a fully functioning publicly-accessible, digital repository for archaeological information with a focused and skilled professional staff. tDAR has a growing number of registered users (6,186 as of 1 September 2014) and content (over 8,000 full-text document files, 17,480 images, and 875 data sets, plus over 360,000 document citation records enhanced and incorporated from the National Archaeological Database). Although most of the information relates to North American archaeology, tDAR includes data from all over the world. Additionally, tDAR is constantly evolving to meet archaeologists' needs in preserving archaeological data—in 2013 tDAR started taking geospatial and scan data and recent 2014 releases addressed tools to increase usability, streamline the curation workflow, and improve data integration tools.

For archaeological data from the US and most international contexts, there is no viable alternative to tDAR as a disciplinary digital repository (Figure 5). At the University of York in England, the Archaeology Data Service (ADS) maintains an archaeological digital repository, but it includes only data from United Kingdom (UK) archaeological contexts or data that are generated by UK researchers. ADS and tDAR do not compete and have partnered on several projects. In the US, the Open Context web-publishing site, which once advertised itself as a digital repository, has substantially modified its services over the last few years. It now focuses squarely on the peer review and publication of archaeological data sets. It uses other organizations to archive data that it publishes. Open Context also requires completely open access to the data it publishes and cannot protect confidential information, e.g. specific site locations, a function that is essential for public agency use.

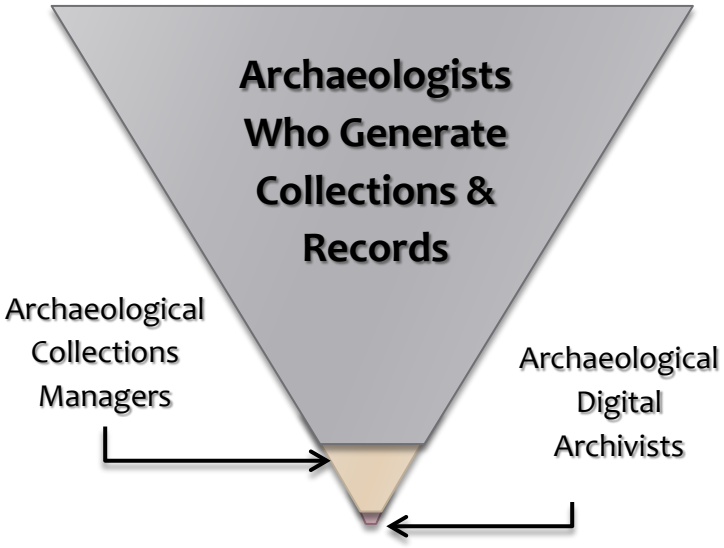

Figure 5: At present, the people who work in the field of archaeology are primarily the ones doing the excavations and writing reports. Far fewer archaeologists are employed in curation facilities caring for the resulting collections, and the only people working specifically to archive for long-term preservation and easier access digital archaeological records generated in the U.S. are at the Center for Digital Antiquity.

There are general-purpose digital repositories, including those operated by universities for data their faculty create or utilize. However, many of these either do not accept or do not

adequately document the data types and metadata categories that archaeologists collect and require for adequate reporting of their results. Because of their general-purpose nature, these repositories cannot offer the functionality that tDAR provides for archaeological data. While they maintain standard technical metadata, they include only very general substantive metadata, seriously limiting both information discovery and reuse. tDAR, on the other hand, allows for the inclusion of detailed substantive metadata specifically tailored for archaeology and for the administrative and management needs of the federal agency. This metadata is essential for data discovery, reuse, and preservation, especially for systematically recorded databases. tDAR structures information and provides a user interface designed for archaeologists and the managers of archaeological information.

tDAR utilizes metadata that conform to standard and widely-used formats including Dublin Core and Metadata Object Description Schema (MODS). Metadata categories are tailored to describe clearly archaeological information and encode spatial, temporal, cultural, material, and other keywords, as well as detailed information regarding authorship, sponsorship, and other sorts of credit that must accompany any use of downloaded data (Appendix B, Part 1). tDAR makes it possible to record full citation information and systematic archaeological and administrative metadata, as well as metadata that is specific to the resource type.

Web-based forms guide data contributors through a streamlined process of metadata entry and file upload. For spreadsheets and databases, this includes documentation of individual data sets, with the ability to map columns to coding sheets and ontologies. tDAR is also able to store and preserve Geospatial files and Sensory Data/ 3D Scan data and the metadata fields are tailored to record important metadata about how the data was produced.

Materials contributed to tDAR can be kept strictly confidential, can be made available to defined lists of individuals, or can be made generally available. Any of these levels of access to the uploaded file can be selected by the individual or organization that creates the tDAR metadata record and uploads the file.

To avoid unnecessary repetitive efforts in metadata entry and maintenance, digital objects in tDAR may be organized into "projects" whose locational, administrative, and other general metadata elements are shared by the project's reports, data sets, images, etc. The tDAR metadata record for each digital object can also "inherit" the project metadata and/or it can be given more specific metadata. In addition, contributors can organize their data into "collections" within tDAR for ease of administration and to more easily control levels of access to the data.

Digital Antiquity curators and technical staff follow practices and procedures for archiving and curating digital files that ensure their long-term preservation and availability for current and future uses. These practices and procedures include:

- daily file backup and protection;
- weekly testing of metadata files and digital files uploaded to detect and remove any file deterioration;
- migration of files into new formats and standards as they develop; and
- secure copies maintained off-site for emergency and disaster protection.

In addition, tDAR metadata and uploaded files (with the exception of files marked "confidential") are indexed by Google and other search engines and are highly discoverable. This function addresses the need to meet open data and access requirements, as well as making the data more accessible and useful to archaeologists performing work at DoD Properties. However, a person's level of access to a file in tDAR is dependent on their user category, whether the resource is a "draft" or "active" resource and whether the file is publicly available. tDAR is therefore able to protect security and confidentiality when required.

There are three general types of tDAR users: unregistered, registered, and contributors. Unregistered users are able to search, find, and view lists of resources in tDAR based on searches that they do of the repository contents. Unregistered users also can view the metadata records, but they are not able to download or view any of the files from tDAR. To download or view actual files, a user needs to register and agree to tDAR's user agreement (see Appendix B, Part 2). Registered users are able to download publicly available documents, i. e., files that have not been marked as "confidential" by the individual or organization who contributed them to tDAR. Registered users also may request access to restricted or embargoed files by contacting the individual or organization who uploaded or authorized the upload of the file. Even registered users cannot access files that contributors or Digital Curators have marked as "draft". Typically, files and metadata records are marked as "draft" while they are part of active projects that Digital Antiquity is working on with clients. For example, the DoD Legacy project tDAR records and uploaded files are marked as "draft" until the review by CRM staffs at the installations is complete. The final type of tDAR user is the contributor. Contributors have to agree to the Contributor agreement (see Appendix B, Part 3) and are able to view and edit resources they have access to, including draft records (see Appendix B, Part 4 for more details). This tiered level of access allows contributors to control access to their materials, and ensures that users are aware that the archaeological records in tDAR should be used appropriately.

tDAR is an open source application developed by Digital Antiquity. Digital Antiquity constantly monitors the use and content of the repository to ensure that newly deposited content is appropriate, it is not infected by any malicious software or users, and performance remains high. The system architecture used for tDAR is designed to scale to growth. In one major episode of adding over 350,000 citation records over a weekend, tDAR did not experience any slowing in system performance. tDAR is set up to take advantage of an economy of scale and has developed pricing models to match this feature. Prices per record,

per file, and per megabyte of storage space needed decrease with increasing the numbers of records, files, and/or megabytes required for any given project.

As a centralized digital archive, tDAR is set up to accept deposits from many individual projects and organizations. Digital Antiquity can provide professional digital archiving services at a much lower cost per deposit than any organization that attempted to provide the same kind of digital archiving for a more limited amount of digital data. Because digital data can be accessed over long distances, there is no practical need that each individual organization include a digital archive with all the services available in tDAR at their home unit(s). Digital data archiving is ideal for an activity that utilizes an economy of scale approach.

Digital Antiquity currently works with public agencies, CRM firms, publishers, research organizations, and individual researchers who are using tDAR to address their archaeological information management needs. Digital Antiquity staff also have reached out to tribal archaeological and historic preservation programs as well. Several have expressed interest in using tDAR, but a lack of funding so far has prevented the development of a project with a tribal program.

One of tDAR's clients is the Bureau of Reclamation (BRec), whose Phoenix Area office is depositing in tDAR technical reports from over 40 years of large archaeological projects done as part of water management projects. The Phoenix area office in concert with Digital Antiquity also is developing links between the archaeological site inventory information in its GIS resource management system and tDAR records related to sites in the inventory. The BRec office also is directing (and funding) CRM firms carrying out current archaeological projects on its behalf to place the digital data generated by these current projects into tDAR as part of their contract responsibilities. The Bureau of Land Management's Permian Basin program in New Mexico is doing the same kind of digital curaton using tDAR.

Digital Antiquity also has completed the first phase of a project with the Air Combat Command of the US Air Force to create digital archives for its bases. Collections for three bases are complete and the project is moving into the second phase, which will include Air Force-wide implementation of the use of tDAR to preserve digital archaeological materials. Digital Antiquity is working with archaeologists and CRM managers in the Air Force to create collections in tDAR for up to 50 different Air Force bases. Digital Antiquity project managers and digital curators will work closely with base CRM and command experts to review the digital documents and data being included in tDAR to ensure that confidential information, mainly specific site locations, as well as any sensitive military information, are shielded from general availability.

Digital Antiquity's successful relationship with federal agencies like BRec, BLM, and the U.S. Air Force's Air Combat Command suggests that other federal agencies might want to use these digital curation services as well. It was therefore chosen for the ECAMDAR Defense Legacy project to see if the DoD as a whole could potentially use tDAR as a resource.

3. Project Description and Objectives

Sara Rivers Cofield

During the proposal stage of the ECAMDAR project, installation PoCs were consulted to seek guidance about how a digital archive could be of benefit to them. A preliminary meeting with project participants and installation PoCs was held on 8 October 2013 at the MAC Lab to discuss the project and solicit questions. While all of the PoCs agreed that their digital documents should be preserved, they did have concerns about costs, security, and information control.

Some of these concerns could be addressed even before the project took place. For example, the question came up as to whether users would be charged fees to access information, such as a monthly subscription. This question could be addressed by existing tDAR policy. Registering for tDAR is completely free. The only charge is the one-time fee for uploading files to the repository. Once a metadata record is added to tDAR, there is no fee for adding to or editing the metadata or uploading replacements of the files. Additionally, viewing the records and searching in tDAR is free for any user and downloading files is free for registered users. This policy is central to Digital Antiquity's mission of making archaeological records as accessible as possible.

Other concerns could only be addressed by trying the tDAR system with actual data. The primary questions this project addresses are as stated above; namely, how could tDAR support the military mission, can tDAR maintain the security standards needed for DoD records, is the program cost effective, and why should the DoD use tDAR instead of caring for their own digital archaeological records. The following project objectives were therefore developed as a framework for addressing these questions.

Question 1: How would partnership with Digital Antiquity using tDAR support the DoD's military mission?

Objectives:

- Describe the relationship between the military mission and cultural resource management laws and mandates.
- Explain the role that digital archaeological records currently have in advancing the military mission.
- Compare the military mission contribution of digital archaeological records that are and are not in tDAR.

Question 2: What are the potential security risks in using tDAR, and can these risks be mitigated?

Objectives:

- Discuss the security considerations inherent in doing archaeology on DoD properties.
- Offer examples of installation policies that prevent archaeology from being a security risk.
- Explain the nature of archaeological documentation as it pertains to potential security breaches.
- Explain tDAR's policies with regard to system security (protection from hackers, malware, etc.)
- Assess tDAR's capabilities regarding redaction and confidentiality access.
- Explain how security concerns were handled for the data included in this project.
- Explain how the techniques used to ensure security in this project may (or may not) be applicable DoD-wide.

Question 3: Is the use of tDAR cost effective?

Objectives:

- Explain the fee structure adopted by tDAR and offer examples of fees associates with different types of projects.
- Consider the ongoing costs to the DoD of establishing its own certified digital archaeological repository.
- Compare the overall costs of using tDAR vs. establishing a DoD digital archaeological repository.

Question 4: Why should the DoD create a partnership with Digital Antiquity and use tDAR instead of its own internal IT resources and staff?

Objectives:

- Survey installation points of contact to establish current methods of managing archaeological data within different DoD settings.
- Evaluate current methods of internal DoD management of digital archaeological data by comparing current practice to DoD mandates, regulations, and guidelines.
- Evaluate tDAR's management of digital archaeological data by comparing current practice to DoD mandates, regulations, and guidelines.
- Point out the strengths and weaknesses of each approach to determine how well each method fulfills DoD requirements for archaeological stewardship.

The objectives outlined above are both ambitious and broad in nature, necessitating an approach that is designed to consider any situation the DoD might encounter in managing its digital archaeological records. Since it would not be realistic to include data from every DoD installation in this evaluation of tDAR, a regional approach was adopted and developed in a manner that would maximize the variables examined. These variables include:

1) **Time:** By using existing data, approximately 20 years' worth of files is included in the project.
2) **Format:** No file formats were excluded from the data sent to Digital Antiquity so as to evaluate how a range of files can be processed and ingested into tDAR.
3) **Archaeological project scale:** Every digital archaeological record associated with existing collections was submitted (e.g. Phase I, Phase II, and Phase III compliance excavations, research excavations, avocational projects, etc.).
4) **Number of files:** The amount of data per project varies from a single record to hundreds of files. This helps establish whether there is a minimum amount of data needed to make submission of a project to tDAR worthwhile.
5) **Number of repositories:** Each archaeological repository has its own system of organization, so it cannot be assumed that data ingestion from one curation facility would work the same way as another. Two facilities therefore contributed to this project to test tDAR's flexibility in accepting data from different curatorial systems.
6) **Number of project archaeologists:** Individual archaeologists and archaeological contract firms can establish their own unique documentation systems as long as they meet professional standards. Thirty different companies or individual archaeologists generated the data included in the ECAMDAR project, which again tests tDAR's flexibility in accepting data.
7) **Number of installations/administrative units:** This study covers 25 installations, and while some of these fall within the same administrative structures (e.g. the Naval District Washington/ NAVFAC Washington includes 16 facilities), many are stand-alone units (e.g. Ft. Meade, Ft, Detrick, Aberdeen Proving Ground, Ft. Lee, etc.).
8) **Different levels of security:** The installations in this study vary from low-security sites such as the Solomons Recreation Center, the USNA Golf Course, and the USNA Dairy Farm, to installations where high-security is needed such as Ft. Detrick, Aberdeen Proving Ground, and the Naval Support Facility Indian Head, which have housed centers for biological weapons, ordnance testing and storage, and nitroglycerin manufacturing, respectively. These facilities therefore cover a broad range that may be representative of the DoD as a whole.

These eight variables allowed rigorous testing of tDAR as a suitable digital repository for the DoD's archaeological data by presenting the project participants with many different scenarios and problems to address. The following sections of this report will discuss the process of ingesting the digital materials into tDAR, the results of these efforts, and the recommendations and guidelines developed along the way.