The Digital Archaeological Record (tDAR)

Notes Before Completing the Application

We have read and understood the notes concerning our application submission.

True

Reviewer Entry

Reviewer 1

Comments:

Reviewer 2

Comments:

CORE TRUSTWORTHY DATA REPOSITORIES REQUIREMENTS

Background & General Guidance

Glossary of Terms

BACKGROUND INFORMATION

Context

R0. Please provide context for your repository.

Repository Type. Select all relevant types from:

Domain or subject-based repository, Research project repository

Reviewer Entry

Reviewer 1

Comments:

Reviewer 2

Comments:
Accept

Brief Description of Repository

The Digital Archaeological Record repository, or tDAR (www.tdar.org), is an international domain repository whose mission is to ensure the long-term preservation of irreplaceable archaeological documents and data and to broadening the access to these resources. tDAR's organizational structure makes the repository highly suitable for preserving archaeological documents and data from new research projects, and existing legacy resources from a variety of cultural heritage projects, all while ensuring this information meets the FAIR principles (findable, accessible, interoperable, and reusable).

Reviewer Entry

Reviewer 1

Comments:
Applicant name updated.

Reviewer 2

Comments:
Accept

Brief Description of the Repository's Designated Community.

As an archaeology domain repository, tDAR is designed to serve the needs of a wide range of archaeologists, researchers, organizations, and institutions who produce, consume, use or manage archaeological and cultural heritage resources. tDAR's Designated Community is global, including Federal, State and Local Government Agencies, Cultural Resource Management and Private Archaeological Consulting Firms, Academic Presses, Academic Researchers, students and educators at all grade levels (https://www.tdar.org/why-tdar/, https://www.tdar.org/using-tdar/).

Reviewer Entry

Reviewer 1

Comments:

Reviewer 2

Comments:
Accept

Level of Curation Performed. Select all relevant types from:

A. Content distributed as deposited, B. Basic curation – e.g. brief checking; addition of basic metadata or documentation, C. Enhanced curation – e.g. conversion to new formats; enhancement of documentation

Reviewer Entry

Reviewer 1

Comments:

Reviewer 2

Comments:
Accept

Comments

The tDAR repository supports two types of deposits, self-service upload and documentation by a user (Type A) and full-service preparation, documentation and upload performed by Digital Antiquity staff members (Types B and C). With Self-Service deposits, users are prompted by detailed, online tDAR metadata web pages and file upload forms to describe and document their information resources. Web Forms are designed to elicit the appropriate metadata fields for each different type of information resource (e.g., data file, document, image, GIS file, etc.). Self-Service metadata is not currently reviewed by DA staff.

With full-service digital curation, the client provides the files (paper or digital) and Digital Antiquity staff members organize the material, digitize paper files, create derivative formats (if needed), create and enter appropriate metadata, upload the files, and prepare redacted versions if needed (https://www.tdar.org/about/policies/contributors-agreement/, https://www.tdar.org/about/policies/accession-policy/, https://tdar-arch.atlassian.net/wiki/spaces/TDAR/pages/557196/Creating+Editing+Resources+in+tDAR).

Reviewer Entry

Reviewer 1

Comments:

Reviewer 2

Comments:
Accept

Insource/Outsource Partners. If applicable, please list them.

Digital Antiquity is an educational and research center hosted by Arizona State University (ASU) and administered by the ASU College of Liberal Arts and Sciences (CLAS; https://thecollege.asu.edu/research/centers) and the School of Human Evolution and Social Change (SHESC; https://shesc.asu.edu/centers). ASU is the legal entity for all services, consultancy and research carried out by DA and provides office space and infrastructure support to allow DA to function. DA’s finances are managed through SHESC and the Office for Research and Sponsored Projects Administration (ORSPA; https://researchadmin.asu.edu/) at ASU. Outsourced partners include:

1. DataCite (https://datacite.org/) – DA purchases DOIs from DataCite and is partnered through the Arizona State University Libraries
2. DataONE (https://www.dataone.org/) – DA is a Tier 1 DataONE member node and tDAR metadata records are exposed through DataONE's platform.
3. Amazon S3/Glacier (https://aws.amazon.com/glacier/) – DA uses S3/Glacier for daily, weekly, and quarterly cloud backups of tDAR files.
4. Arizona State University Core Facilities house DA's virtualized cloud servers (https://cores.research.asu.edu/research-computing/about).
5. Arizona State University Library (https://lib.asu.edu/) – Digital Antiquity has a partnership and Memorandum of Agreement with the ASU Libraries such that the Library will maintain, ensure access to, and provide long-term preservation of the digital objects and related metadata currently in the tDAR repository if Digital Antiquity should cease to exist. The current 5-year agreement is in effect until July 2027.

Reviewer Entry

Reviewer 1

Comments:

Reviewer 2

Comments:
Accept

Summary of Significant Changes Since Last Application (if applicable).

Since submission of the previous application, all reviewer comments were addressed. One major update is a renewed agreement between the Arizona State Libraries and the Center for Digital Antiquity such that the Library will maintain, ensure access to, and provide long-term preservation of the digital objects and related metadata currently in the tDAR repository if Digital Antiquity should cease to exist. The new agreement runs till 2027 with a provision to be renewed upon review.

Other changes include the organization of suggested passages to different sections, clarification between the Center for Digital Antiquity and their governance of the tDAR repository, clarification of how tDAR verifies archived materials, and an expanded discussion regarding the storage of data.

For this most recent revision (28/10/2022), we 1) clarified the applicant in the Background sections, 2) in Section 7 changed the compliance level to 4 and added text at the end of the section explaining the integrity of the resources and user submissions, 3) in Section 8 we added paragraph at the end of the section explaining how resources submitted to the repository in the future (likely 2023) via the self-service function will receive a review from digital curators, 4) in Section 10 condensed narrative as to not completely repeat our online "Terms of Use" and added links where appropriate, and 5) in Section 14 changed compliance level from 3 to 4 as reviewer 2 felt we had met the criteria with this narrative.

Reviewer Entry

Reviewer 1

Comments:

Reviewer 2

Comments:

Other Relevant Information.

The creation of Digital Antiquity and tDAR involved ASU faculty and staff at the School of Human Evolution and Social Change (SHESC), the ASU Library, and ASU Research Computing. Digital Antiquity staff and administrators continue to maintain close working relationships with these ASU academic units. In addition, Digital Antiquity seeks advice from, and is guided by, a Board of Directors (https://live-digant.ws.asu.edu/board-of-directors/). The Board includes experts in finance, law, not-for-profit organization management, digital data management, cultural resource management, and experts from other universities. DA has established formal partnerships with the Society for American Archaeology (SAA; https://www.saa.org/), the Society for Historical Archaeology (SHA; https://sha.org/) and the American Cultural Resources Associate (ACRA; https://www.acra-crm.org/). DA has a strong partnership with the Archaeological Data Service (ADS; https://archaeologydataservice.ac.uk/), housed in the University of York, and collaborated with ADS (https://archaeologydataservice.ac.uk/) to publish and update the Guides to Good Practice series (https://guides.archaeologydataservice.ac.uk/g2gp/Main). DA is a member of the ARIADNEplus project (https://ariadne-infrastructure.eu/), a network of European Union and other organizations promoting heritage data preservation. DA is also involved in the European Cooperation in Science and Technology (COST) SEADDA project (Saving European Archaeology from the Digital Dark Age; https://www.seadda.eu/), which aims to bring a global community of archaeologists and data management specialists together to share expertise, and create resources that allow them to address digital preservation problems in their home countries.

Reviewer Entry

Reviewer 1

Comments:

Reviewer 2

Comments:
Accept

ORGANIZATIONAL INFRASTRUCTURE

1. Mission/Scope

R1. The repository has an explicit mission to provide access to and preserve data in its domain.

Compliance Level:

4 – The guideline has been fully implemented in the repository

Reviewer Entry

Reviewer 1

Comments:
4 – The guideline has been fully implemented in the repository

Reviewer 2

Comments:
4 – The guideline has been fully implemented in the repository
Accept

Response:

tDAR (https://www.tdar.org/) is an international domain repository whose mission is to ensure the long-term preservation (https://www.tdar.org/about/) of irreplaceable archaeological data and to broadening the access (https://www.tdar.org/why-tdar/access/) to these resources (including documents, data sets, images, GIS files, 3D scans and other resource types). The materials preserved in tDAR document the archaeological record throughout the world, the efforts of the archaeological and scientific community, and the material and social characteristics of the cultures studied. tDAR is operated and maintained by the applicant, the Center for Digital Antiquity.

Digital Antiquity, through the development and maintenance of the tDAR repository, supports broad access to archaeological data. Browsing or searching the tDAR repository enables users to identify digital documents, data sets, images, and other kinds of archaeological data for a number of uses, including research, learning, and teaching. tDAR enables users to download data files while maintaining the confidentiality of legally protected information and the privacy of digital resources on which a researcher is still working. Our intelligent searching tools make it easy to search within any or all of the fields containing the rich metadata associated with each record. General searches also search within the text of each publicly accessible file. Users are able to narrow their search by resource type, or spatially by using the map tool. Additionally, tDAR's metadata is indexed by major search engines, ensuring that the information within the repository can be located by a wide audience.

tDAR and its parent organization, the Center for Digital Antiquity, are dedicated to ensuring the long-term preservation of digital archaeological data. These data document the archaeological record, the efforts of the archaeological and scientific community, and the material and social characteristics of the cultures studied. Digital data are fragile. Both the file formats and the media on which they are stored are in danger of becoming obsolete over time, making the archaeological data they contain irretrievable. Records archived in tDAR are preserved for access and use today and will be maintained into the distant future by:

● Regularly and systematically checking the files in the tDAR repository to ensure that no deterioration has occurred (Section VII).

● If file deterioration is detected, taking steps to remedy it.

- Periodically migrating and/or refreshing the digital files to provide for their long-term accessibility and preservation.
- Planning for obsolete technology.
- Maintaining files in open and preferable formats, and accommodating new industry standards for archaeological information.
- Storing rich, descriptive metadata for all resources.

Reviewer Entry

Reviewer 1

Comments:

Reviewer 2

Comments:

2. Licenses

R2. The repository maintains all applicable licenses covering data access and use and monitors compliance.

Compliance Level:

4 – The guideline has been fully implemented in the repository

Reviewer Entry

Reviewer 1

Comments:
4 – The guideline has been fully implemented in the repository

Reviewer 2

Comments:
4 – The guideline has been fully implemented in the repository
Accept

Response:

Arizona State University is the legal entity for all services, consultancy and other contracts carried out by Digital Antiquity, with an archive level deposit license for use with both individual and institutional depositors that has been implemented to protect the rights of depositors, users and repository.

Registered tDAR users are required to accept the Terms of Use Agreement (https://www.tdar.org/about/policies/terms-of-use/) and, if contributing, must also accept the Contributor’s Agreement (https://www.tdar.org/about/policies/contributors-agreement/). While users do not have to be registered to browse public metadata, they must be registered in order to contribute data or download files. These agreements are based on the Creative Commons Attribution 3.0 Unported License (https://creativecommons.org/licenses/by/3.0/).

All registered users must accept the Terms of Use Agreement when creating a free, tDAR account. In accepting this, users agree they will copy or distribute content under the following conditions:
1. Users must accompany all uses and applications of this content with proper citation and attribution (as provided on the tDAR metadata page).
2. For any redistribution of tDAR content, users must clearly include proper citation and attribution information and make clear to others the license terms of this work.
3. Users must not use tDAR content in ways that could be reasonably expected to lead, directly or indirectly, to damage to the archaeological record.
4. Users acknowledge that neither The Center for Digital Antiquity nor its sponsors and associates guarantee the accuracy or usability of the content and further agree that they may not hold any of these parties liable for any direct or consequential damage arising from their use of tDAR or its content.
5. Users are responsible for ensuring that their use of tDAR and its content is consistent with applicable law.

When registering, users also have the opportunity to select whether they wish to contribute resources and/or metadata to tDAR. This prompts the user to accept the Contributor's Agreement form. A contributor is defined as, "the person or institution responsible for the files and metadata contributed to tDAR." This agreement covers the following Sections:
A. Authority, Appropriateness, and Accuracy
B. License, Copyright, and Re-use
C. Confidential and Embargoed Information
D. Users' Terms of Use
E. Unacceptable Content
F. Resolution of Issues Concerning tDAR Content

As noted in Section E of the Contributor's Agreement, Resolution of Issues Concerning tDAR Content (https://www.tdar.org/about/policies/contributors-agreement/), a series of steps are followed by Digital Antiquity in order to identify issues of compliance and the appropriate response. These steps include:
1. Issues concerning content may arise either through internal reviews or from external reports that Digital Antiquity receives;
2. Digital Antiquity will review all such issues raised and that during such review, files and associated metadata may be withdrawn from public access;
3. In attempting to resolve issues, Digital Antiquity will attempt to solicit comment both from the reporter of the issue and from the contributor, using the contributor's most recent email address on file with Digital Antiquity;
4. Having reviewed the issue, Digital Antiquity will, at its sole discretion, determine the appropriate resolution which may include the removal or redaction of materials from tDAR; and,
5. Under no circumstances will Digital Antiquity refund any deposit fee or assume any cost or liability incurred by the contributor related to contributing or distributing the contested information.

This section in the Contributor’s Agreement is supported by the Center’s Accession Policy (https://www.tdar.org/about/policies/accession-policy/) which states, “Digital Antiquity retains the right to review and remove files and metadata records that do not comply with its policies.”

Reviewer Entry

Reviewer 1

Comments:

Reviewer 2

Comments:

3. Continuity of access

R3. The repository has a continuity plan to ensure ongoing access to and preservation of its holdings.

Compliance Level:

4 – The guideline has been fully implemented in the repository

Reviewer Entry

Reviewer 1

Comments:
4 – The guideline has been fully implemented in the repository

Reviewer 2

Comments:
4 – The guideline has been fully implemented in the repository
Accept

Response:

Digital Antiquity is dedicated to ensuring the long-term preservation of the digital data stored in the tDAR repository. As described on the tDAR Preservation web page, all data in the tDAR repository are stored and protected in perpetuity (https://www.tdar.org/why-tdar/preservation/, https://www.tdar.org/about/policies/preservation-and-curation-policy/).

Digital Antiquity manages the tDAR repository’s development, maintenance, and financial and technical sustainability. Digital Antiquity is an academic and research center affiliated with the School of Human Evolution and Social Change in the College of Liberal Arts and Sciences at Arizona State University. Those units provide administrative, and some financial support, as well as review and guidance for both Digital Antiquity and the tDAR repository.

Digital Antiquity undergoes quarterly reviews by its Board of Directors every year, an annual review by the Director of SHESC, and is subject to a thorough review by CLAS every 5 years. Each type of review examines staffing, performance, business operations and financial stability. Technical, business and domain experts are involved in each of these reviews.

Following the OAIS reference model (Reference Model for an Open Archival Information System; http://www.oais.info/) Digital Antiquity designed the tDAR repository with technical procedures in place to ensure that digital files and metadata are curated and preserved for the medium and long term (see answer R15).

Digital Antiquity and the Arizona State University Library have a formal Memorandum of Agreement stating that in the event that Digital Antiquity is no longer able to support the development and maintenance of the tDAR repository, all files in the repository will be transferred to, and curated by, the ASU Digital Repository (https://repository.asu.edu/). Upon receiving notice from Digital Antiquity that it is unable to continue to maintain tDAR, staff of the ASU Libraries have agreed to provide assistance with the smooth transfer of digital objects and related metadata from tDAR to the Libraries digital repository. In this case, the Libraries will maintain, ensure the access to, and provide long-term preservation of the digital objects and related metadata as they exist at the time of transfer from tDAR as part of its repository. For tDAR records that include digital objects marked “confidential,” the Libraries will include the metadata in its repository and put the confidential digital file(s) in a “dark archive” that will ensure long-term preservation and the possibility of retrieval when appropriate. Content contributors will be notified and provided with the opportunity to download copies of their files. The agreement is open to renegotiation and renewal on a 5-year basis; the current agreement is in effect from 2022 to 2027.

Reviewer Entry

Reviewer 1

Comments:

Reviewer 2

Comments:

4. Confidentiality/Ethics

R4. The repository ensures, to the extent possible, that data are created, curated, accessed, and used in compliance with disciplinary and ethical norms.

Compliance Level:

4 – The guideline has been fully implemented in the repository

Reviewer Entry

Reviewer 1

Comments:
4 – The guideline has been fully implemented in the repository

Reviewer 2

Comments:
4 – The guideline has been fully implemented in the repository
Accept

Response:

tDAR and its parent organization, Digital Antiquity, are dedicated to ensuring the long-term preservation of digital archaeological data. These data document the archaeological record, the efforts of the archaeological and scientific community, and the material and social characteristics of the cultures studied. Because many of the files archived in tDAR contain both culturally sensitive information and site location information, tDAR has several features that assist users in complying with ethical and legal requirements.

To begin, Digital Antiquity policies and help pages provide guidance for contributors on confidentiality and ethics. Contributors are pointed to the Guides to Good Practice (https://guides.archaeologydataservice.ac.uk/g2gpwiki/) website for a discussion of these issues.

In general, the archaeological data files in the tDAR repository do not typically contain personally identifiable information (PII). Upon registering to use tDAR, users are asked for first and last name, email, organization, affiliation/interest, and work phone (optional) (https://core.tdar.org/account/new). tDAR users have the means to manage their own personally identifiable information from the tDAR user profile. Users can control what information Digital Antiquity has access to, such as phone numbers or contact information. Digital Antiquity will not sell users' contact information and will not disclose it to third parties unless legally obligated to do so. If a user wishes to have personally identifiable information removed from tDAR they are directed to send an email requesting such. Digital Antiquity will not remove your name from public documents or tDAR records, but will ensure that all contact information within the application is removed. Requests for the removal of personally identifiable information within specific files in tDAR, such as PDF documents, should be directed to the authors or contributors of those materials.

These policies are outlined to registrant when asked to confirm that that have read and accept the tDAR User Agreement (https://www.tdar.org/about/policies/), which address tDAR's Privacy Policy (https://www.tdar.org/about/policies/privacy-policy/) and Terms of Use (https://www.tdar.org/about/policies/terms-of-use/), and Contributors Agreement (https://www.tdar.org/about/policies/contributors-agreement/).

Digital Antiquity does collect other information to provide better service to all of our users. Information is collected actively as users register, browse, or interact with tDAR and as users interact more directly with Digital Antiquity staff. The information we collect includes:
1. Contact Information: We use contact information to notify users of requests within the system (such as contact or download requests for confidential files), to assist in the troubleshooting, or other issues. Contact information is not shared with third parties unless that information is required to provide a service such as processing credit card information or

sending email.
2. Device information: For example, what web browser you use or if you are using a mobile device.
3. Log Information: What a user does on tDAR is captured in server logs including the IP address of your computer. Our log data tracks searches, downloads, pages viewed, and information that may be useful for debugging purposes when users save resources or perform other actions within the system. This information helps us understand how users interact with tDAR and assists us in troubleshooting issues within the system. Log data is stored for two weeks and then deleted automatically.
4. Usage Information: We track searches, downloads, page views, and other usage information independent of personal information.
5. Cookies and other Identifiers: tDAR uses cookies to assist with authentication and storage of user preferences.
6. Use of Third-Party Software and Services: Digital Antiquity uses third party software to provide some services such as web analytics (e.g., analytics performed by Google), hosting of fonts or other files, or credit card processing.

This information is used to provide, maintain, protect, and improve tDAR and assist in developing new features and services.
tDAR’s mission is to make data discoverable, accessible, and reusable; however, in the United States (and other countries), federal and state laws protect cultural heritage data (e.g., National Historic Preservation Act; https://www.nps.gov/subjects/historicpreservation/national-historic-preservation-act.htm and the Archaeological Resources Protection Act; (https://www.nps.gov/archeology/tools/laws/arpa.htm), particularly with regard to the specific location of archaeological sites (to prevent looting) and the display of culturally sensitive material (such as mortuary imagery). tDAR was designed to meet these confidentiality and ethical standards and has policies and automated processes in place to ensure compliance. In some cases, the contributor agrees to make a reasonable effort to designate as “confidential” any contributed file that would be reasonably expected to endanger in situ archaeological sites if it were made publicly available. The contributor assumes any liability for improper disclosure of information that contributor should have, but did not, designate as confidential.

Legally protected site locations are obfuscated in tDAR's geographical search and map display. All site locations shown on the maps on the tDAR web site are obfuscated https://www.digitalantiquity.org/wp-uploads/2017/05/20170503-DA-Security-Summary-Final.pdf, https://tdar-arch.atlassian.net/wiki/spaces/TDAR/pages/557165/Confidentiality+and+Security). As stated in our security document:

“The metadata describing tDAR records allow contributors to designate UTM coordinates or specific site locations on a map. If these spatial designations are smaller than one square mile, the tDAR software will obfuscate the spatial data when displaying this information to users who have not logged in or have not been explicitly granted access. When obfuscated, spatial designations will be randomized and will display an area greater than one square mile.”

Contributors have the option to upload materials to tDAR but embargo public access to those resources for up to 5 years. The metadata for each record is always “Public” but the file is marked as embargoed, with a stated availability date (https://tdar-arch.atlassian.net/wiki/spaces/TDAR/pages/557219/Managing+Security+and+Access).

In response to the adoption of the European Commission General Data Protection Regulations (https://ec.europa.eu/info/law/law-topic/data-protection_en) (GDPR) Digital Antiquity updated the tDAR Contributor's Agreement (https://www.tdar.org/about/policies/contributors-agreement/). In addition to these changes, it is standard procedure to prohibit the upload of files to tDAR that contain Social Security numbers, credit card numbers, driver license numbers, or other categories of protected information. Documents or other data that contain this kind of information will be marked as "draft" until the resource owner redacts and reuploads the file.

While tDAR does not request confirmation that data collection or creation was carried out in accordance with legal and ethical criteria prevailing in the data producer's geographical location or discipline, files uploaded by Digital Antiquity staff are reviewed for sensitive cultural information (e.g., mortuary related) and explicit references to exact site locations (e.g., maps, UTM coordinates, etc.). If files contain such information, Digital Antiquity staff members contact the contributor and discuss options for site location obfuscation and/or redaction of data within each file. In the case of redaction, the full, original version of the file is stored in tDAR but marked confidential. The redacted copy is the public version of the file that is available for viewing and download.

Furthermore, contributors have the option to redact confidential or sensitive information, and Digital Antiquity staff discusses this option with contributors. Contributors also have the ability to limit access to files using either an embargo or by marking their files Confidential (Private). Metadata for confidential files is public, but the file itself cannot be downloaded without permission. A contact email is provided for users to inquiry about gaining access to the confidential file(s). Decisions to grant access (or not) are made by the data contributors/creators/owners (https://tdar-arch.atlassian.net/wiki/spaces/TDAR/pages/557227/Planning+Your+Data+Contribution, https://tdar-arch.atlassian.net/wiki/spaces/TDAR/pages/557219/Managing+Security+and+Access, https://tdar-arch.atlassian.net/wiki/spaces/TDAR/pages/557165/Confidentiality+and+Security, https://tdar-arch.atlassian.net/wiki/spaces/TDAR/pages/557185/Access+Rights+in+tDAR). The Terms of Use (https://www.tdar.org/about/policies/terms-of-use/) and Contributor's Agreement (https://www.tdar.org/about/policies/) outline the ethical norms for using tDAR data or adding data to the repository. If not compliant with these policies, Section E of the Contributor Agreement outlines the procedures in place to handle non-compliance (https://www.tdar.org/about/policies/contributors-agreement/).

Reviewer Entry

Reviewer 1

Comments:

Reviewer 2

Comments:

5. Organizational infrastructure

R5. The repository has adequate funding and sufficient numbers of qualified staff managed through a clear system of governance to effectively carry out the mission.

Compliance Level:

4 – The guideline has been fully implemented in the repository

Reviewer Entry

Reviewer 1

Comments:
4 – The guideline has been fully implemented in the repository

Reviewer 2

Comments:
4 – The guideline has been fully implemented in the repository
Accept

Response:

The tDAR repository is managed by the Center of Digital Antiquity, a research unit within Arizona State University (ASU), the School of Human Origin and Social Change (SHESC), and the College of Liberal Arts and Sciences (CLAS). Staff working for tDAR are employed by ASU and the School of Human Evolution and Social Change (SHESC; formerly the Department of Anthropology). This hierarchical structure of reporting and affiliation is appropriate for the tDAR community of users, and provides day-to-day continuity and stability.

In addition to the relationship with SHESC, the Center for Digital Antiquity has a 5-year Memorandum of Agreement (MOA) with the Arizona State University Library. This MOA ensures access to, and long-term preservation of, the digital objects and related metadata in the tDAR digital repository in the event that Digital Antiquity is unable to continue to support the development and maintenance of tDAR. The agreement is open to renegotiation and renewal on a 5-year basis; the current agreement is in effect from 2022 to 2027 and is available upon request.

Financial support for Digital Antiquity derives from several sources including: institutional support from Arizona State University; grant funding; and revenue from full-service digital curation and preservation services provided by Center staff members (https://www.tdar.org/about/history/). ASU and SHESC provide support for basic overhead needs including office space, utilities, computer network and security, and administrative and business support. The Digital Antiquity business model supports the tDAR repository and Center staff through depositor payments (both self-service and full-service deposit). Charges for self-service uploads are collected by credit card through existing ASU business infrastructure (Recharge Center; https://researchadmin.asu.edu/recharge-centers). Business contracts with full-service clients are negotiated by DA administrators and then administered through the ASU business office. Grant funding provides an additional revenue stream.

Digital Antiquity staff members (https://live-digant.ws.asu.edu/our-team/) include six full-time employees, with additional (temporary) student employees and post-docs. Current staffing includes an executive director, a program manager of digital preservation, and program manager of business operations, a programmer, a digital curator, and an administrative assistant. The Center also employs a variable number of graduate and undergraduate student digital curators (2-4 typically) and post-docs each year. All Digital Antiquity staff members have educational backgrounds, training and work experience appropriate to their position.

o Director –an archaeologist who manages the Center (projects, staff, grants, research), participates in grant-funded projects, and engages in sales and marketing with prospective clients.

o Program Manager for Business Operations –with a business background whose focus is sales and marketing, client services, and representing Digital Antiquity at conferences and other venues.

o Program Manager for Digital Preservation - an archaeologist who performs day-to-day project management of digital curators and students, participates in grant-funded projects and manages outside client projects, focused on data management and metadata. They also teach data management workshops offered by Digital Antiquity to clients and at various conferences.

o Programmer – an IT professional whose responsibilities include maintaining the core tDAR repository software, systems administration and website development.

o Digital Curator –with experience in metadata entry and data management, primarily focused on fulfilling full-service client contracts.

o Administrative Assistant – the staff member is the business and office manager, with duties including accounting, web site content updates and social media posting.

Digital Antiquity staff attend local, national, and international conferences and workshops focused on archaeology, heritage management, computer applications in archaeology, and digital archiving annually. In addition, continuing education classes in archaeology, data management, and repository technology are available through various departments at ASU. Staff members are encouraged to complete relevant classes and provided with time-release during work hours to attend relevant educational opportunities. Annual Employee Performance Evaluations contain specific suggestions for staff development activities (including attendance at conferences, workshops, and enrolling in classes) tailored to individual staff members' responsibilities and background.

Finally, regarding organizational infrastructure, DA is overseen by a Board of Directors (https://live-digant.ws.asu.edu/board-of-directors/). The Board includes experts in finance, law, not-for-profit organization management, digital data management, cultural resource management, and information technology, plus individuals from the University of Arkansas, Arizona State University, the SRI Foundation, the Bureau of Land Management, University College Dublin, and the University of York.

Reviewer Entry

Reviewer 1

Comments:

Reviewer 2

Comments:

6. Expert guidance

R6. The repository adopts mechanism(s) to secure ongoing expert guidance and feedback (either inhouse or external, including scientific guidance, if relevant).

Compliance Level:

4 – The guideline has been fully implemented in the repository

Reviewer Entry

Reviewer 1

Comments:
4 – The guideline has been fully implemented in the repository

Reviewer 2

Comments:
4 – The guideline has been fully implemented in the repository
Accept

Response:

Digital Antiquity works closely with a network of individuals, organizations and agencies within both its subject area and the wider data management communities to remain relevant to the community it serves. To begin, DA is governed by an independent Board of Directors and as a research center at Arizona State University, Digital Antiquity maintains close collaborative ties with faculty, students and staff at ASU. Feedback from those communities serves as both advisory and as a point-of-contact with Designated Community users.

For the purposes of expert guidance and feedback, the Digital Antiquity Board of Directors (https://live-digant.ws.asu.edu/board-of-directors/) includes experts in finance, law, not-for-profit organization management, digital data management, cultural resource management, and information technology, plus individuals from the University of Arkansas, Arizona State University, the SRI Foundation, the Bureau of Land Management, University College Dublin, and the University of York.

Each year, Digital Antiquity hosts quarterly Board of Directors conference calls, as well as one 2-day face-to-face Board meeting in Tempe, Arizona. The DA Director reports both to the Board of Directors and to the Director of the ASU School of Human Evolution and Social Change. Agenda items for Board meetings include advice on financial stability and general operations, staffing, tDAR technical infrastructure, sales and marketing, and the tDAR user experience.

Additional external advisors are consulted as needed, for example, to design and implement tDAR website upgrades, to improve digital preservation procedures, or to improve the tDAR repository infrastructure. As an example of the successful interaction with our external advisors, DA, in collaboration with the UK-based Archeological Data Service (https://archaeologydataservice.ac.uk/), have worked together to create and develop a series of digital and print-based Guides to Good Practice (https://guides.archaeologydataservice.ac.uk/g2gpwiki/) which provide subject-specific guidance to data producers on appropriate data management and preservation strategies within the archaeological and historic environment sectors. These guides have been researched and written by a mixture of experts and practitioners within the heritage community, data management specialists and DA staff. These volumes are ‘active’ publications and kept under periodic review from the wider profession, and, where necessary, are updated to legislate for these comments. In instances where new technologies are deployed, or techniques developed, DA will actively engage with experts and data technologists to create new guidance so that good data standards can be assured across its designated community. All Guides to Good Practice are actively promoted to the profession, and form the basis of the DAs’ own in-house Data Procedures.

tDAR’s internal advisors consists of various faculty, students, and staff members at Arizona State University including faculty in relevant academic departments (e.g., computer science, digital humanities, archaeology), the ASU Libraries, the University Technology Office and others with expertise in cultural heritage, technology, and digital security.

The tDAR Contact Us web page (https://www.tdar.org/about/contact-us/) includes an email address, mailing address, and phone and fax numbers for users to directly communicate with DA staff. Finally, there is a mechanism in tDAR that allows users to contact resource owners to request access or to suggest a correction to metadata. This function is located on each resource metadata page under the Download section. If a user would like to report a technical issue or problem with tDAR website they can send a note to comments@tdar.org.

The tDAR website also has links soliciting feedback from users at the bottom of every page. Social media sites including Twitter (https://twitter.com/DigArcRec), Instagram (https://www.instagram.com/digitalantiquity/), and Facebook (https://www.facebook.com/digarcrec) provide avenues for Designated Community feedback. Staff members write a Digital Antiquity newsletter, which is sent via email to users. Back issues of the eNewsletter are available on the Digital Antiquity website (https://www.digitalantiquity.org/enews/).

Digital Antiquity curators and technical staff actively partake in professional workshops and conferences (local, national, and international) keep abreast of changes in file format obsolescence, best practices in digital preservation, and changing needs of the archaeology and cultural heritage user community. In addition, Digital Antiquity staff members attend several regional, national, and international conferences, workshops and meetings where Designated Community members can talk with staff about issues related to Digital Antiquity. DA staff routinely participate at The Society for American Archaeology annual meeting (https://www.saa.org/annual-meeting), the Society for Historical Archaeology annual meeting (https://sha.org/conferences/), the Archaeological Institute of America annual meeting (https://www.archaeological.org/programs/professionals/annual-meeting/), the Arizona Historic Preservation Conference (https://www.azpreservation.org/conference), and the Plains Anthropological Conference (https://plainsanthropologicalsociety.org/annual-meeting), and the Southeast Archaeology conference (https://www.southeasternarchaeology.org/). Digital Antiquity staff members are partners in a number of international grant-funded projects including ARIADNEplus (https://ariadne-infrastructure.eu/) and SEADDA (https://www.seadda.eu/). These, along with the Computer Applications and Quantitative Methods in Archaeology conference (https://caa-international.org/) serve to connect Digital Antiquity staff members with a broad, international user base.

Reviewer Entry

Reviewer 1

Comments:

Reviewer 2

Comments:

DIGITAL OBJECT MANAGEMENT

7. Data integrity and authenticity

R7. The repository guarantees the integrity and authenticity of the data.

Compliance Level:

4 – The guideline has been fully implemented in the repository

Reviewer Entry

Reviewer 1

Comments:
3 – The repository is in the implementation phase

Reviewer 2

Comments:
4 – The guideline has been fully implemented in the repository
Accept

Response:

tDAR maintains a bespoke Java-based Collections Management System (CMS) and Object Management System (OMS) which facilitates the documentation of all datasets and collections. On ingest, all files are stored in their original format and an MD5 checksum (Message-Digest algorithm 5) is created and stored (https://tdar-arch.atlassian.net/wiki/spaces/DEV/pages/720950/tDAR+technical+overview#tDARtechnicaloverview-FixityChecksandFileIntegrity). Digital Antiquity staff perform fixity checking (comparing the MD5 signature of a file on ingest with the current file in the filestore) on all materials in the tDAR filestore on a weekly basis. The weekly timeframe was chosen because it is one-half of the total time of the local, on-site backups (i.e., backups are done every 2 weeks) which ensures that any file can be restored. If an MD5 checksum does not match, the system sends an email to the lead tDAR developer for review. After analysis by the developer, a corrupted file would be replaced with a backup version that does match the MD5 checksum (https://tdar-arch.atlassian.net/wiki/spaces/DEV/pages/720950/tDAR+technical+overview#tDARtechnicaloverview-FixityChecksandFileIntegrity).

tDAR has two options for depositing resources: self-service and curated. Files that are uploaded using the full-service digital curation option are delivered to DA staff (either in paper format or digital format), and processed with a standard workflow. With the self-service option, contributors are required to fill in metadata for the following fields: title, year created, description, and a contact name and email, for each record that is uploaded. Standard tDAR website metadata entry forms ensure compliance and help normalize metadata entries, including suggesting standard terms. tDAR provides extensive help text on the metadata entry pages and via the “Upload Toolkit” website (https://www.tdar.org/using-tdar/upload-toolkit/) and the tDAR wiki (https://www.tdar.org/using-tdar/upload-toolkit/).

Regarding data completeness, tDAR does not ensure that the submitter’s data is complete for self-service. tDAR accepts each resource “as is.” However, tDAR does have a program which verifies that the resource is valid for each file type, but not data completeness.
There are effectively two forms of verification. The first is filename validation. When a user submits a resource metadata upload form, tDAR has a program that verifies that the file extension belongs to the list of accepted extensions for that resource (e.g., documents may be DOC, DOCX, TXT, or PDF). Second, after passing the file extension check, tDAR attempts to create derivatives of the file (for example, producing a thumbnail preview of a document or an image). If a file is invalid or corrupted, tDAR will still accept the file but will notify the user of any errors that occurred while attempting to create these derivatives. For curated projects, tDAR’s digital curators visual check data for completeness and note any issues with the data in the metadata documentation. Under a proposed future pricing model, DA intends to include curator time for doing content checks and QA/QC for self-service uploads.

There are two ways in which data are changed in tDAR. First, files are migrated to preservation formats automatically to ensure long-term access. Second, occurs when a file is replaced by a user with a newer version. In both cases, the system has a resource revision log and there is a log entry describing the change in the resource. The system creates an XML export of the resource metadata and saves that as well. Regarding metadata changes, any entry added to the Resource Revision Log results in the system creating an XML export of the resource metadata and saves the new information.

When a resource owner uploads a new version of a resource, tDAR retains the old one with a new name (suffix) with a version number associated with the name. For every change made to the resource and metadata a new XML file is generated documenting those changes; changes are tracked with log files that are only viewable by tDAR administrators. Log files show the revision history of each resource, along with a link to the previous version of the resource.

For international metadata standards, tDAR is based on Dublin Core (https://dublincore.org/) concepts, augmented with custom fields for archaeology. tDAR is guided and abides by standard data conventions but do not enforce named schema/standards for archaeological data (https://tdar-arch.atlassian.net/wiki/spaces/DEV/pages/720900/tDAR+Document+Metadata+Mapping+to+MODS+and+DC). We also use the OAIS Model's Submission Information Package (SIP), Dissemination Information Package (DIP), and Archival Information Package (AIP) (https://tdar-arch.atlassian.net/wiki/spaces/DEV/pages/720950/tDAR+technical+overview).

While DA does not strictly adhere to the OAIS information model, there are rough equivalents to the SIP, AIP, and DIP concepts.

● (SIP): The files that the user includes in their submission are initially placed in what we call a user's "personal filestore" in BagIt Packaging Format. Once the user submits their resource metadata form, tDAR processes the submitted files and removes these files from the personal filestore.

● (AIP) After processing, the archival files from the user are placed in the tDAR filestore along with resource metadata in XML format.

● (DIP) Derivatives (thumbnails, etc.) are also located in the tDAR filestore, however, they are not included in offsite backups.

More information about the tDAR Filestore is available at https://tdar-arch.atlassian.net/wiki/spaces/DEV/pages/720950/tDAR+technical+overview#tDARtechnicaloverview-ThetDARFilestore.

In regards to data changes, tDAR does not have a strategy for tracking changes apart from keeping versions of files. Additionally, tDAR does not make changes to user datasets, but has a feature that allows for any public user to contact the resource owner if there is an issue with a dataset. In the future, if tDAR were to house dynamic, living datasets, users will be notified if their data submission into the repository requires any changes.

tDAR asks for detailed information and contact details for all data creators, copyright holders, publishers, funders, etc. for a resource to provide provenance for data and allow the appropriate audit trail to be documented and maintained, at least from the point of accession into the repository. Such metadata is created by those most familiar with the dataset, typically the data creator or depositor. This includes information that can be used to identify individuals or organizations (hosting institution, ORCID ID, etc.); this personal information is gathered by the system with the users' informed consent as outlined in the tDAR Privacy Policy (https://www.tdar.org/about/policies/privacy-policy/). Metadata for both collections and digital objects is added to allow a full audit trail for resources to be preserved.

tDAR maintains links to different resources in several ways. First, all resources within a project or collection are linked to one another. If a user accesses a single resource they are then able to return back to the original project/collection page to view other associated items in that collection. Second, datasets that have associated coding sheets and ontologies are linked to the original dataset. Finally, resource creators are linked to all of the resources they have uploaded into tDAR, regardless if the resources are related.

The integrity of resources are ensured through the process detailed above, however, due to the nature of our workflows and user submissions, complete authenticity is not possible. Currently, data and resources uploaded to tDAR are not reviewed for authenticity. CTS defines authenticity as “the degree of reliability of the original deposited data and its provenance, including the relationship between the original data and that disseminated, and whether or not existing relationships between datasets and/or metadata are maintained.” DA does not verify the reliability of the original data with the resource owner, assuming that the resource owner has an ethical obligation to submit “authentic” resources which is a key component in what a contributors agrees to in the contributors agreement (https://www.tdar.org/about/policies/contributors-agreement).

Reviewer Entry

Reviewer 1

Comments:
In the absence of any verification at the point of deposit a level 3 compliance level is recommended.

Reviewer 2

Comments:

8. Appraisal

R8. The repository accepts data and metadata based on defined criteria to ensure relevance and understandability for data users.

Compliance Level:

3 – The repository is in the implementation phase

Reviewer Entry

Reviewer 1

Comments:
3 – The repository is in the implementation phase

Reviewer 2

Comments:
3 – The repository is in the implementation phase
Accept

Response:

As a domain repository, Digital Antiquity accepts data and metadata related to a wide range of archaeological and cultural heritage investigations. Selection is guided by the tDAR Accession Policy (https://www.tdar.org/about/policies/accession-policy/), as well as the Contributor’s Agreement Policy (https://www.tdar.org/about/policies/contributors-agreement/), and the Terms of Use Policy (https://www.tdar.org/about/policies/terms-of-use/).

Digital Antiquity does not accept non-archaeological data in tDAR, unless the data are closely related to an archaeological project or of general relevance to the discipline and to the Designated Community. Examples might include tree ring data (related to dating archaeological sites) or paleoclimate data. Non-archaeological datasets may also be accepted as part of a package of materials related to a larger archaeological research project (e.g., files describing the geology of a project region). The relevance of these non-archaeological data sets is evaluated by Digital Antiquity staff members on ingest, as stated in the Accession policy and the Contributor's Agreement policy. As stated in the Accession policy "Digital Antiquity retains the right to review and remove files and metadata records that do not comply with its policies."

As noted in question R0, the tDAR repository supports two types of deposit, self-service upload by a user and full-service upload performed by Digital Antiquity staff members. With Self-Service deposits, users are prompted by online tDAR metadata web pages and file upload forms to fully describe and document their information resources. There are three mandatory metadata fields in tDAR: title, year, and abstract/description. Ideally, these three metadata fields allow users to properly use the resource. Without entries in those fields, tDAR will not allow the resource to be saved in the repository. tDAR does not have an explicit procedure for determining the academic sufficiency of self-service metadata beyond the help menus and support documents (https://tdar-arch.atlassian.net/wiki/spaces/TDAR/pages/557227/Planning+Your+Data+Contribution).

With full-service digital curation Digital Antiquity staff members create comprehensive metadata for all information resources, and solicit additional documentation as needed from the information resource owner.

tDAR employs a modified metadata schema based in Dublin Core with additional structured metadata elements appropriate for archaeology and cultural heritage information resources. Additionally, the tDAR metadata schema has been mapped to both Dublin Core and the MODS metadata schema which allows for metadata mapping through our OAI-PMH server. (https://tdar-arch.atlassian.net/wiki/spaces/TDAR/pages/557196/Creating+Editing+Resources+in+tDAR, https://tdar-arch.atlassian.net/wiki/spaces/DEV/pages/720900/tDAR+Document+Metadata+Mapping+to+MODS+and+DC).

Digital Antiquity staff members have access to all files and metadata stored within tDAR. Staff may, on behalf of contributors, access and review materials to ensure that materials are being properly preserved and curated. Digital Antiquity retains the right to review and remove files and metadata records that do not comply with its policies (https://www.tdar.org/about/policies/preservation-and-curation-policy/). Furthermore, tDAR enforces the submission of at least three metadata values before a user can submit their resource for archiving.

tDAR is designed to accept file formats most commonly used by our Designated Community, which currently stands at 25 separate file formats. A list of file formats that tDAR accepts is available on tDAR (https://core.tdar.org/contribute). tDAR has checks in place to ensure that the resource owners can only submit the preferred file formats. If a depositor attempts to submit a file that is not one of the preferred formats, tDAR will not allow the file to be uploaded. If the file type is not listed in the field "tDAR Supported File Types" then depositors must convert their files prior to submission.

The tDAR file upload function is designed to automatically reject file types that are not supported. The contributor receives an error message ("file extension is not valid for the file X") explaining that the file type is not accepted in tDAR. Digital Antiquity staff will work with Contributors to assist them in converting files to an accepted format before submission.

As stated in the Accession Policy (https://www.tdar.org/about/policies/accession-policy/), Digital Antiquity retains the right to review and remove files and metadata records that do not comply with its policies. Staff members have the authority to access and review files in the repository to ensure that the material is appropriate to the Center's mission, and that the files are properly preserved and curated.

In addition, files and associated metadata may be withdrawn from public access without removing them from the tDAR file store. Digital Antiquity staff will then attempt to resolve the issue, soliciting information both from the reporter of the issue and from the contributor, using the contributor's most recent email address on file with Digital Antiquity. Having reviewed the issue, staff will determine the appropriate resolution which may include the removal or redaction of the files in question from tDAR.

DA is working to implement new workflows that will ensure resources submitted by self-service or user submitted, will receive some level of review by curation staff. This review will focus primarily on metadata completeness and standards and work with user to ensure corrections are completed before the resource can become Active in the repository. While not at extensive as the full-curation service, this additional curation review will ensure that all resources are meeting the same metadata criteria for completeness and relevance. We hope to have a workplan within the next year, to then start this process by 2024.

Reviewer Entry

Reviewer 1

Comments:

Reviewer 2

Comments:

9. Documented storage procedures

R9. The repository applies documented processes and procedures in managing archival storage of the data.

Compliance Level:

4 – The guideline has been fully implemented in the repository

Reviewer Entry

Reviewer 1

Comments:
4 – The guideline has been fully implemented in the repository

Reviewer 2

Comments:
4 – The guideline has been fully implemented in the repository
Accept

Response:

All data and metadata (AIP, DIP, and SIPs) are stored and backed up by Digital Antiquity staff on the Arizona State University network and two other storage locations. Uploaded files immediately go into the filestore on a server housed at Knowledge Enterprise (KE) Research Computing and Data Service (https://cores.research.asu.edu/computing-and-data-services). Daily backups of these new resources, and metadata, are uploaded to Amazon's east coast Glacier Server (AWS) (https://aws.amazon.com/glacier/). Every quarter the system also creates a copy of the complete filestore (files and metadata), which is also stored in Amazon Glacier. The programming staff backs up the originally submitted files as well as the resource metadata. We generally do not include derivatives such as thumbnails in our backups. Combined, these daily and quarterly backups allow for the recovery of tDAR in the case of catastrophic failures on one or more servers. Furthermore, the programming staff makes copies on external hard drives quarterly. A copy of the tDAR metadata database, housed in a Postgres database, is zipped and sent to Amazon Glacier daily and quarterly. All of the data is encrypted using GNU Privacy Guard (GPG).

Besides the benefit of having redundant copies of a backup, multiple backup venues have trade-offs that influence which backup we may choose when restoring files.

- ASU Backups: ASU Research Computing (which host's tDAR's hardware) performs their own backups, and is the most convenient choice in terms of requesting a retrieval of a backup, however, their coverage only extends a couple weeks.
- External Drive backups: we maintain these in order to have a recovery option that is more convenient (and less expensive) than retrieving files via Amazon Glacier.
- Amazon Glacier backups: we store our files in what Amazon refers to as their "deep archive" class of storage. It is very durable and low-cost, however, retrieval requests may take 24-48 hours to initiate and would take several days to complete. We anticipate that we would only use this recovery option in the event of a catastrophic event that takes out our datacenter and our external backups.

There are several ways in which processes and procedures are managed. All processes within tDAR are documented in the online documentation program (or wiki) known as "Confluence" (https://tdar-arch.atlassian.net/wiki/spaces/DEV/overview). These pages detail aspects of development configuration, database scheme, source code structure, integration testing, deploying tDAR on a server, and Eclipse setup and configuration. These developer pages are maintained by the tDAR programming staff as needed.

tDAR's file storage and management model is heavily influenced by the California Digital Library's Micro-Services model (http://ijdc.net/index.php/ijdc/article/viewFile/154/217/). Data is stored on the file-system in a pre-determined structure described as a PairTree filestore (https://confluence.ucop.edu/display/Curation/PairTree). The filestore maintains versions of all of the data and metadata in tDAR. The core structure of the filestore is a set of nested directories which parallel the unique ID of metadata stored within the Postgres database. The pair-tree also partition data on the file-system into manageable chunks. Each branch of the filestore is a folder for each record. Data associated with each tDAR record is stored in a structure inspired by the D-Flat convention (https://confluence.ucop.edu/display/Curation/D-flat) ensuring a consistent organization of the record (https://tdar-arch.atlassian.net/wiki/spaces/DEV/pages/720950/tDAR+technical+overview#tDARtechnicaloverview-ThetDARFilestore).

Digital Antiquity attempts to create an archival version that maintains the data fidelity, appearance, and functionality of all materials submitted to tDAR. A checksum MD5 procedure is completed and stored in Postgres database for each file uploaded into tDAR. When the file is archived the Filestore another Checksum MD5 is completed and the MD5 strings are compared to ensure they match. If there is a mismatch, the files are flagged and included in a list of unmatched MD5s, which is sent to the tDAR programming staff. For tDAR AWS backup service, all files are zipped, added to the AWS and AWS performs another checksum MD5; this can be compared to the original zip file sent to AWS.

tDAR works in partnership with ASU KE and Amazon for physical storage space; these groups are responsible for checking the deterioration of storage media. All procedures for checking storage media are documented by KE (https://cores.research.asu.edu/sites/default/files/Research-Computing-Facilities-Statement_June2020.pdf).

Reviewer Entry

Reviewer 1

Comments:

Reviewer 2

Comments:

10. Preservation plan

R10. The repository assumes responsibility for long-term preservation and manages this function in a planned and documented way.

Compliance Level:

4 – The guideline has been fully implemented in the repository

Reviewer Entry

Reviewer 1

Comments:
4 – The guideline has been fully implemented in the repository

Reviewer 2

Comments:

4 – The guideline has been fully implemented in the repository
Accept

Response:

Digital Antiquity, the university center charged with overseeing the tDAR repository, has a stated mission to extend our knowledge of the human past and improve the management of our cultural heritage by permanently preserving digital archaeological data and supporting their discovery, access, and reuse. This is accomplished through the use of tDAR, the digital repository that houses digital resources from archaeological investigations. Digital Antiquity also promotes good practice in the use of digital data, provides technical advice to the heritage community, and supports the deployment of digital technologies.

Digital Antiquity's Preservation and Curation Policy operationalizes this mission (https://www.tdar.org/about/policies/preservation-and-curation-policy/). This preservation and curation policy is divided into the following process:

1. Future Access to Information
2. Staff Access to Information in tDAR
3. Modification of Metadata
4. Programmatic Access to and Modification of Files by Digital Antiquity
5. Creation of Archival Copies
6. Fidelity of Archival Copies
7. Dissolution of Archive

All digital objects in tDAR are stored in the original format in which they were submitted and in a preservation format. The original format is maintained at a bit level. The preservation format is selected to maintain a high level of stability though long-term migration. To maintain a high level of usability, a derivative format may also be created to ensure compatibility with contemporary software. Finally, tDAR will accommodate new industry standards as they evolve. For example, Microsoft Excel files (.xls) are maintained as submitted, but are transformed to ASCII CSV (comma separated value files) as a preservation format. In addition, tDAR will migrate Excel files to future versions (e.g., .xlsx) for dissemination. Regarding the preservation of data semantics, in addition to the general metadata that tDAR uses to describe resources tDAR allows for additional metadata specific to datasets. When supported by the underlying file format (e.g. Microsoft Excel, Microsoft Access), tDAR will retain information about the tables within a dataset, as well as column names and data-types. Users can further define table and/or column metadata using the Dataset Column Metadata edit pages (link to section). Furthermore, tDAR allows users to download the original, archival version of a dataset file. In the case of downloading an original dataset file, tDAR does not remove any metadata associated with the file, rather it allows users to download a "translated" version of the dataset in CSV format. Due to the limitations of the CSV file format, these translated datasets will not retain any of the column or table-level metadata that may have been included in the original/archival version of a dataset.

Digital Antiquity will also attempt to create an archival version that maintains the data fidelity, appearance, and functionality of all materials submitted to tDAR. Barring the ability to produce an archival version with the exact appearance and functionality, Digital Antiquity will attempt to ensure that the content (e.g., text in a document, tabular data within a data set) is preserved.

tDAR maintains a Contributors Agreement which outlines the rules that contributors must agree to when they deposit files in tDAR (https://www.tdar.org/about/policies/contributors-agreement/). In general, tDAR supports an “open access” approach to sharing of archaeological information. Accordingly, the contributor recognizes that the files and metadata contributed to tDAR can be shared with tDAR users. There are, however, certain situations in which information should be treated as confidential or otherwise restricted (see Contributor Agreement https://www.tdar.org/about/policies/contributors-agreement/). For example, tDAR has the ability to allow a contributor to specify levels (public, confidential, and embargoed) of access to files, while all metadata are publicly available (https://tdar-arch.atlassian.net/wiki/spaces/TDAR/pages/557219/Managing+Security+and+Access).

tDAR also features an Upload Toolkit that discusses submission information and archival standards for long-term preservation (https://www.tdar.org/using-tdar/upload-toolkit/ . Submission information standards includes details on:

1. Preparing to upload data with tDAR
2. Resources Types accepted
3. Upload Data Requirements
4. Why Metadata are Essential
5. Tools for Organizing Your Archive
6. Organizing Your Data for Archiving
7. Developing a Metadata Style Guide
8. Benefits for Future Research
9. tDAR Templates

Regarding measures to ensure proper archival actions are taken, Digital Antiquity provides two methods to upload into the tDAR repository; self-service and full-service deposit. For self-service deposit, it is up to the contributor to ensure that they are meeting the submission guidelines outlined within tDAR’s policies. Self-service users can refer to the tDAR Upload Toolkit for guidance on selecting resources to archive (https://www.tdar.org/using-tdar/upload-toolkit/). Digital Antiquity is currently undertaking a plan where future self-service submissions will undergo quality control and assurance measures made possible as part of an accession fee that provides funding for digital curator time.

In the case where contributors are working with Digital Antiquity digital curators for full-service deposits, the digital curators guide the contributors through the process of selecting the appropriate resources to upload. Contributors often encounter either time or monetary constraints, and the digital curators can guide them through the process of selecting resources that best meet their archival needs. Furthermore, tDAR, in conjunction with the Archaeology Data Service (ADS), have created a Guides to Good Practice (https://guides.archaeologydataservice.ac.uk/g2gpwiki/) which guides contributor decisions on what materials to archive.

Reviewer Entry

Reviewer 1

Comments:

Reviewer 2

Comments:

11. Data quality

R11. The repository has appropriate expertise to address technical data and metadata quality and ensures that sufficient information is available for end users to make quality-related evaluations.

Compliance Level:

3 – The repository is in the implementation phase

Reviewer Entry

Reviewer 1

Comments:
3 – The repository is in the implementation phase

Reviewer 2

Comments:
3 – The repository is in the implementation phase
Accept

Response:

There are several approaches to addressing data and metadata quality in tDAR. First, resources and their associated metadata are deposited and curated in one of two ways: by contributors themselves (self-service) or by trained digital curators on staff at Digital Antiquity (full-service). While the majority of the curation is done under the “full-service” option, there are methods to ensure the quality of the data uploaded by self-service contributors. These include:

1. File verification that identifies corrupt files
2. File verification that identifies files types that are not accepted by tDAR
3. Metadata entry assistance; guidance on required fields and other documentation

Apart from the system checks and fields that assist data entry, Digital Antiquity offers a range of tutorials, guides, workshops, and recommendations for contributors in order to ensure the quality of data for their self-service deposits:

1. Workshops with professional organizations and societies
2. Guides to Good Practice (https://guides.archaeologydataservice.ac.uk/g2gpwiki/)
3. Style guides (https://www.tdar.org/using-tdar/upload-toolkit/#developingAMetadataStyleGuide)
4. Help documentation (https://www.tdar.org/using-tdar/)

For each file uploaded, contributors are required to fill in metadata for the following fields: title, year created, description, and a contact, if the file is denoted as confidential or embargoed. In addition to required metadata fields, during entry, fields are accompanied with side text to help the user understand each field and enter the appropriate information. Proposed future directions include charging a basic accessioning fee for quality control and assurance on all self-service uploads to be done by the center's digital curators.

The full-service option for metadata entry involves the center's digital curators working with clients to upload their material with activities ranging from creating guides to working to archive the files from start to finish. By offering the full-service option, we are able to ensure that the files and data entered are properly handled and correctly curated.

While tDAR does require metadata for each resource to enhance its reuse, tDAR does not have quality control checks to ensure the completeness and understandability of self-service materials deposited. tDAR operates under the assumptions that it is the responsibility of the contributor to meet these criteria, and that they understand that tDAR is an online archive that ensures the long-term preservation, access, and reuse of the materials they deposit. For metadata, tDAR does have a function that requires a contributor to input a minimum number of metadata fields, ensuring that resources are minimally documented for reuse. If a contributor does not add the minimum metadata fields, they are alerted to this fact during the submission process and are unable to proceed with the submission.

Proposed future directions include charging a modest accessioning fee for quality control and assurance checks on all self-service uploads to be done by the center's digital curators. While full curation will not be done by the Center's curators, they will be able to identify issues and send feedback that must be dealt with before the resources become Active and searchable in tDAR.

Links to citation indices for resources uploaded into tDAR are not provided; however, DOI's assigned to resources are indexed and identified in Google Scholar and are linked to a contributor's Google Scholar profile automatically.

Reviewer Entry

Reviewer 1

Comments:

Reviewer 2

Comments:

12. Workflows

R12. Archiving takes place according to defined workflows from ingest to dissemination.

Compliance Level:

4 – The guideline has been fully implemented in the repository

Reviewer Entry

Reviewer 1

Comments:
4 – The guideline has been fully implemented in the repository

Reviewer 2

Comments:
4 – The guideline has been fully implemented in the repository
Accept

Response:

Digital Antiquity accepts a wide range of archaeological data and archaeologically-related data in tDAR. Materials contributed to tDAR include digital files (e.g., documents, data sets, images, scans, and other types of digital files) related to a wide range of archaeological investigations and topics (e.g., field studies of various scales and intensities, archives and collections, and historical, methodological, synthetic, or theoretical studies) and related metadata.

Digital Antiquity staff, along with the DA and tDAR websites, communicate with, and to, users regarding how their files are handled. With in-person communications to full and self-service clients, knowledgeable staff, including the Director, Program Managers, Curators, and Programmers, are able to explain how tDAR archives and makes materials accessible. Additionally, the tDAR website has multiple Help features that detail how the files are handled (https://www.tdar.org/using-tdar/; https://tdar-arch.atlassian.net/wiki/home).

To ensure these materials are properly archived, Digital Antiquity follows a workflow for consultation, planning, and implementing archiving for full-service clients regarding 1) planning and consultation, 2) data ingest (OAIS SIP) and review (human and automated), 3) archiving (OAIS AIP), and 4) access and dissemination (OAIS DIP). For full-service clients the process follows these general steps outlined below:

Planning and Consultation: this step is handled by the Center Director and or Program Managers, who offer guidance on starting an archival project, answering questions about 1) pricing and invoicing, 2) assistance with deciding which data files to archive (selection), 3) communicating file formats and their sustainability, 4) assessing hard-copy digitization needs, along with conversion to PDF/a and conversion tips, and 5) discussion on metadata standardization, controlled vocabularies, and consistency (metadata style guide).
Data Ingest (SIP): this step is handled by our Program Manager and Digital Curators. tDAR accepts a wide variety of file

formats, that are specified on the tDAR website (https://core.tdar.org/resource/add). Generally, the type of file does not impact the overall workflow, however, some files, such as images, if not properly documented, may take longer to assign adequate metadata. Likewise, for documents, redacting site locational information, both in the form of maps and within the text, may take longer depending upon document length.

During this stage, DA curators work on 1) file redaction and confidentiality, 2) metadata standardization, 3) collection and sub-collection creation (for logical organization and to improve discoverability). There are also automated steps which include 1) creation a log file for each information resource, 2) creation of a MD5 hash (checksum) for each resource (file fixity), and 3) site location obfuscation.

Data Archiving (AIP): the digital curators are responsible for the following steps: 1) creating collection and sub-collection creation (for logical organization and to improve discoverability), 2) file organization, 3) metadata creation and input, and 3) redaction of sensitive information. Automated steps include: 1) storage of the original files, 2) creation and storage of backup copies of the original file in new locations, 3) creation of derivative(s) in archival format, and 4) the creation and storage of backup copies of the derivative file(s) in new locations.
Data Dissemination (DIP):

● Assign DOIs
● Set resource appropriate permission(s) (i.e., public, embargoed, secured)
● Customer options to create custom collections and/or search features

This workflow is flexible depending upon client/user needs, but follows this general pattern to ensure that the files are properly archived. And to assist in this workflow, specifically for full-service projects, digital curators keep track of their progress using a project tracker; specifically, a spreadsheet that documents completed tasks for each resource.

Regarding workflow roles, Program Managers, Digital Curators, and Programming Staff can perform the following duties after files have been archived in tDAR (as needed):
1. Access all files and metadata stored within tDAR. Staff can, on behalf of contributors, access and review materials to ensure that materials are being properly preserved and curated. For example, for full-service archiving resources, such as PDFs, go through optical character recognition and converted to PDF/A, as part of the standard workflow for documents.
2. Perform minor edits on all materials within the repository. These changes may include minor spelling and grammatical edits or normalization of keyword or common search terms.
3. Produce derivative copies of the files submitted. Through the submission process the tDAR application will, among other actions, open all files submitted, extract text and other information, and create derivative files such as thumbnails.
4. Create archival versions of files that maintain the data fidelity, appearance, and functionality of all materials submitted to tDAR. Barring the ability to produce an archival version with the exact appearance and functionality, Digital Antiquity will attempt to ensure that the content (e.g., text in a document, tabular data within a data set) is preserved.

tDAR has different levels of security depending upon the content of the files uploaded. tDAR does not house files or data with personally identifiable information, or information that requires Institutional Review Board approval on human subjects. Security is typically related to either culturally sensitive information (e.g., images of Native American burials) and site location coordinates. For both of these, the contributor has the ability to assign permission to particular individuals and at the same time secure the files so they cannot be downloaded by the public without contributor approval. These security measures are part of the tDAR workflow, whether for self- or full-service contributors.

Regarding quality checking of outputs, with full-service curation we check the quality of files to ensure files are properly archived. Self-service outputs do not undergo the same review. Likewise, there are not quantitative checks on outputs/resources.

The self-service curation workflow is almost entirely completed by the contributing user. As suggested in the name, the actions that are handled by DA staff within the full-service curation model are to be handled by the user and are not checked by curators. While the automated tasks are still in place, all other functions are the responsibility of the user. These tasks include:

- Planning: the user must identify appropriate file formats and select what files can be uploaded. They will then have to calculate how much space is necessary and the corresponding price using our pricing calculator (https://core.tdar.org/cart/add). Digitization of physical resources must be completed at this point.
- Data Ingest: once the files are gathered by the user and the billing account established in tDAR, a user will then upload the appropriate files, add metadata, organize into collections, and assign file and access restrictions as necessary.
- Review: manual review of the organization and metadata is left up to the user.
- Access and Dissemination: once the files are uploaded and made Active, a DOI is automatically created and are now searchable within the repository. The user can now share and handle permissions.

To assist self-service users, Digital Antiquity offers a range of tutorials, guides, workshops, and recommendations for contributors as mentioned in R11. Made specifically with self-service users in mind, is the Upload Toolkit which goes through each stage described above and offers best practices and links to more in depth help documentation found in the Using tDAR section of the site (https://www.tdar.org/using-tdar/upload-toolkit/).

Any issues or problems with these workflows are tracked in our issue tracking management system, Jira. Jira is a ticket software that allows DA staff to track and collaborate on problems encountered in our workflow or the tDAR program associated with that workflow. Jira also facilitates communication with DA staff and tDAR users to ensure timely and appropriate responses to user issues and comments.

Reviewer Entry

Reviewer 1

Comments:

Reviewer 2

Comments:

13. Data discovery and identification

R13. The repository enables users to discover the data and refer to them in a persistent way through proper citation.

Compliance Level:

4 – The guideline has been fully implemented in the repository

Reviewer Entry

Reviewer 1

Comments:
4 – The guideline has been fully implemented in the repository

Reviewer 2

Comments:
4 – The guideline has been fully implemented in the repository
Accept

Response:

tDAR provides enhanced discovery, search and retrieval of archaeological data both via directly from the tDAR website and internet search engines (e.g., Google). tDAR's content is indexed by all of the major search engines. In addition, browsing, basic search, and advanced search tools are available from the tDAR website. Users can search within any or all of the custom metadata fields associated with each record. Files in tDAR are full-text indexed, thus supporting searches for any text string in a file. Spatial search tools using PostgreSQL with PostGIS extension provide another avenue for discovery, as users can draw a polygon around an area using the map tool. This returns a results list of all resources that are within or that intersect the search box.

tDAR provides a search box on every webpage. The advanced search feature has a dedicated webpage (https://core.tdar.org/search) and "Help" documentation for the different search options is also available online (https://www.tdar.org/using-tdar/searching-tdar/).

To facilitate discovery, tDAR has a searchable metadata catalog. However, because archaeological descriptions and culture histories vary from country to country and continent to continent, there is no single, international metadata standard for archaeology. As a domain repository, tDAR uses standard archaeological concepts and controlled vocabularies to describe resources; our metadata categories are easily mapped to Dublin Core and MODS standards.

Another important method for facilitating findability is through the use of persistent identifiers, thus tDAR mints a unique DOI (using DataCite; https://datacite.org/) for each information resource. The DOI is shown as a link on each resource page in tDAR, and is included on the default cover page that is added to each file that is downloaded or exported from tDAR.

tDAR does facilitate machine harvesting of metadata by exposing Dublin Core, Extended Dublin Core, and MODS metadata via an OAI-PMH server running on the main tDAR server. tDAR also facilitates machine harvesting of metadata via DataCite's search site (https://commons.datacite.org/) and through the ARIADNEplus Portal (https://ariadne-infrastructure.eu/portal/).

tDAR's metadata is indexed by Google and other web search engines facilitating discovery and access to the data in the repository. We also expose Google Scholar tags in the HTML page of tDAR records to allow metadata to be collected via Google Scholar.

tDAR is listed in the following registries and agencies/organization websites as a recommended repository:
1. Registry of Research Data Repositories (re3data; https://www.re3data.org/)
2. Elsevier Database linking (https://www.elsevier.com/authors/author-resources/research-data/data-base-linking)
3. Open Access Directory maintained by the School of Library and Information Science at Simmons College (http://oad.simmons.edu/oadwiki/Data_repositories)
4. DataOne (https://old.dataone.org/news/dataone-welcomes-tdar)
5. U.S. National Science Foundation, Social, Behavioral and Economic Sciences (SBE) (https://www.nsf.gov/funding/pgm_summ.jsp?pims_id=11690)
6. American Cultural Resources Association (https://www.digitalantiquity.org/wp-uploads/2019/03/2019-DPeter-et-al-Digital-Data-Curation-and-Access-Why-Be-Involved-ACRA-reformatted-FPM.pdf)
7. American Anthropological Association (https://www.americananthro.org/StayInformed/Content.aspx?ItemNumber=1914)
8. Society for American Archaeology (https://www.saa.org/membership/member-benefits), and
9. Society for Historical Archaeology (https://sha.org/announcements/sha-conference-abstracts-now-available-on-tdar/)

tDAR does include a recommended citation for each resource in the repository. The Citation information is prominently listed on the main metadata page for each information resource (file) in tDAR and is included on the default cover page that is added to each document that is downloaded or exported from tDAR. In the future, we intend to create a function that allows users to export the citation to their preferred bibliographic software.

Reviewer Entry

Reviewer 1

Comments:

Reviewer 2

Comments:

14. Data reuse

R14. The repository enables reuse of the data over time, ensuring that appropriate metadata are available to support the understanding and use of the data.

Compliance Level:

4 – The guideline has been fully implemented in the repository

Reviewer Entry

Reviewer 1

Comments:
4 – The guideline has been fully implemented in the repository

Reviewer 2

Comments:
4 – The guideline has been fully implemented in the repository
Accept

Response:

To ensure resources can be reused by the archaeological community, the metadata requested for inclusion by tDAR users are structured or standardized descriptive information, related to archaeological information, that documents the resources in the archive (https://www.tdar.org/using-tdar/upload-toolkit/#whyMetadataAreEssential). tDAR's metadata provides essential context for archival resources, such as authors, dates, geographical location, cultural affiliation, file types, and contacts (https://www.tdar.org/wp-content/uploads/2020/06/Project.docx). It can also include identifiers to help track data submission, access rights, and critical information about how the data relate to other elements of the archive. The metadata that users enter when they upload data to tDAR are stored with the data files so that future users will always have access to that critical contextual information about the resource.

Types of Metadata
tDAR's metadata (https://www.tdar.org/wp-content/uploads/2020/06/Project.docx) is loosely grouped into several categories (described below). These types of metadata are collected at different levels (e.g. project or individual resource) but particular elements may be recorded for data from the project level down to the specific file level. In addition, certain metadata standards may record elements of metadata that function at a number of levels (e.g. 'Author' may aid resource discovery as well as provide administrative information). The three categories relevant to archaeological projects are project, resource, and file metadata.

Project Metadata is largely recorded at a broad level for an entire project/archive irrespective of the techniques used and covers elements such as period terms or dates, site and artifact keywords, project details, site codes and geographic location. Often, much of this information is included within documents in the archive (e.g., site reports). This level of metadata is designed to allow the user to create a comprehensive description and allows for easy retrieval of datasets or documents, and is more about the project and its results. The Dublin Core standard is a good example of a metadata standard which incorporates a number of descriptive and resource discovery focused elements.

Resource Metadata refers to the “middle tier” of metadata elements that are relevant to the individual resources that make up an archival project. A resource, by this definition, may include one or multiple files, depending on the context. A simple example might be a PDF document, image, or access dataset. A more complex example might be a sensory data scan, a GIS shapefile, or even some databases where a single functional 'resource' consists of multiple files.

The purpose of this metadata largely overlaps with the project level, with a descriptive or research discovery emphasis, but focuses on details that typically vary within the context of a larger project. These may include specific bibliographic citation fields, dates, locations, material types, or keyword-list elements that do not crosscut all resources in a project. In a sense, these metadata are more about the results reported in the individual resource rather than in the overarching project, and are documented primarily to aid discovery.

File Metadata is specific and applied at the level of individual files. File metadata incorporates information on hardware and software along with validation methods such as checksums. In many cases, if the data is to be deposited in a digital archive, it is the archive itself that will generate much of this metadata. However, the data creator must often supply a number of elements, often during the process of data creation.

A fourth broad category, Administrative Metadata, exists within all of the above and covers elements such as creation and acquisition as well as alteration and version control. Included within this metadata is information concerning intellectual property rights. Such information can be recorded at a general level (e.g., ownership for an entire dataset may be held by one person or organization) but should also be recorded for specific techniques or datasets where personnel or authors and intellectual property right holders differ. Administrative metadata in tDAR also extends to the user who uploads the resource. All tDAR users must register an account and create a billing account to deposit materials. All of this information is tied to a collection or resource and provides an additional set of metadata, though not publicly available.

The metadata created in tDAR allows users to quickly and easily identify available resources and connects them to those that they need. However, for this to work effectively, the metadata has to be implemented accurately and in a standard format. A commonly used format for project-level metadata is Dublin Core (https://dublincore.org/). Within tDAR, the basic Dublin Core metadata schema has been extended to include archaeology-specific fields as well as to include fields from the Metadata Object Description Schema (MODS) (https://www.loc.gov/standards/mods/, https://tdar-arch.atlassian.net/wiki/spaces/DEV/pages/720900/tDAR+Document+Metadata+Mapping+to+MODS+and+DC).

Element Description
1. Basic Information Basic metadata includes association of a resource with a larger project, lifecycle status (e.g., draft, active, deleted), Language, Year Created, Abstract/Description and Physical Storage Location.
2. Bibliographic Metadata Specific to documents, uses MODS metadata fields to describe resources.

3. Resource Creators These fields are used to properly credit individuals and institutions for their contribution to the resource. “Role” refers to the individual’s main role for the resource (e.g., creator, editor, etc.).
4. Resource Specific or Agency Identifiers Describes agency or project identifiers used (e.g., Smithsonian Trinomial, AZ State Museum—ASM—number).
5. Investigation Type(s) Lists all investigation types relevant to the resource (e.g., Research Design, Site Monitoring, Data Recovery/Excavation).
6. Site Description Information Includes Site Name, Site Type (controlled vocabulary), and keywords
7. Material Type(s) Artifact types covered by the resource (e.g., Ceramic, Fauna, Metal, Dating Sample).
8. Cultural Term(s) Includes Culture (e.g., PaleoIndian, Archaic) and user created values.
9. Temporal Coverage Includes Temporal Terms (e.g., Pueblo IV), Coverage Dates, Date Type (e.g., Calendar, Radiocarbon),
Start/End Years and Description (e.g., Calibrated, Range of ± N years).
10. General Keyword(s) User-created values to describe aspects of the resource not covered by other metadata.
11. Spatial Terms Includes Geographic Terms (e.g., Death Valley), Coordinates
12. Resource Provider The institution that authorized the resource to be archived and disseminated.
13. Individual and Institutional Roles Name, Current Email Address, Institution and Role of those involved with the resource.

Archaeology specific metadata are the key to making data searchable in tDAR. When an archaeologist searches tDAR they often use keywords terms particular to their sub-discipline (e.g., cultural affiliation, site type, artifact type), thus the creation of these metadata at the onset of archiving assists others in retrieving similar types of information.

To ensure that data resources are understandable and reusable, tDAR requires that some metadata to be added to a resource before it can be archived and made publicly available. There are three metadata fields that are required to make the resource understandable and reusable; title, year, abstract/description. As stated previously, Administrative Metadata are required as it is tied to a user account/profile. Though not displayed to the public, this metadata can assist Digital Antiquity Staff should any questions arise about a resource.

In addition to requiring metadata to be submitted, contributors are encouraged to include the submission of Coding Sheets (https://www.tdar.org/using-tdar/creating-and-editing-resources/coding-sheets-create-and-edit/). Coding Sheets may contain information that allows users to decode the coded values for the columns in a spreadsheet or database. It is common practice to provide a text-based Coding Sheet with published datasets so that users can interpret data and replicate your methods. Coding Sheets allows tDAR to make sense of the columns in a dataset, thus a user does not need to necessarily “translate” the meaning of the coded values manually when viewing a spreadsheet or database. The Coding Sheet can also define any codes used in data collection, and ontologies, which define the terminology used.

The files archived in tDAR are in formats broadly used by archaeologists, primarily in the United States, and internationally. Archaeologists archive various kinds of documents including:
1. Reports of archaeological field investigations, articles and presentations, field or lab notes, catalogs, dissertations or

theses, collections and historical research, and historical documents and correspondence about archaeological resources, research projects, and organizations;
2. Spreadsheets, databases, and coding sheets that describe archaeological data sets about artifacts, features, sites, or other archaeological phenomenon;
3. Photographs, maps, and illustrations of archaeological resources or related to archaeological investigations; and,
4. Data about archaeological resources collected by various sensors, e.g., GPS, GIS, Resistivity, GPR, and various sonar instruments.

Based on this type of information, tDAR accepts the following file formats (https://core.tdar.org/contribute):
1. Documents
a. PDF Documents (.pdf)
b. Microsoft Word (.doc, .docx)
c. Rich Text Documents (.rtf)
d. Plain Text Documents (.txt)
2. Datasets
a. Comma Separated Values (.csv)
b. Tab Separated Values (.tab)
c. Microsoft Excel (.xls, .xlsx)
d. Microsoft Access (.accdb, .mdb)
3. Images
a. Tagged Image File Format (.tiff, .tif)
b. JPEG Image (.jpg, .jpeg)
c. Portable Network Graphics (.png)
d. Other (.bmp, .gif, .pict)
4. Geospatial Data
a. Shapefiles
b. Geodatabases
c. Georectified images (GeoTIFF & GeoJPG)
5. Virtual
a. Remote Sensing Files
b. 3D Scans (OBJ & E57)

All of these formats are anticipated to be available long-term and are the current industry standards. In the future, if the broader archaeological community begins to adopt new file formats, tDAR will consider adopting those new file types. On an annual basis, Digital Antiquity will review new file formats and consider their appropriateness for inclusion into tDAR. For example, HEIC raw image files created from Apple cameras may be considered as they are becoming more ubiquitous.

In addition to general metadata about individual resources, additional table metadata for datasets can be added. For example, each dataset uploaded undergoes character recognition, where the column type is identified (Uncoded, Coded, Measurement, Count, or Filename). Furthermore, users can edit the name displayed for the column to help users understand the columns contents, category or subcategory that best describes the data, search visibility (Visible, Confidential, or Hidden), add a narrative of the column description, and if available, map the data to an existing ontology.

Reviewer Entry

Reviewer 1

Comments:

Reviewer 2

Comments:

TECHNOLOGY

15. Technical infrastructure

R15. The repository functions on well-supported operating systems and other core infrastructural software and is using hardware and software technologies appropriate to the services it provides to its Designated Community.

Compliance Level:

4 – The guideline has been fully implemented in the repository

Reviewer Entry

Reviewer 1

Comments:
4 – The guideline has been fully implemented in the repository

Reviewer 2

Comments:
4 – The guideline has been fully implemented in the repository
Accept

Response:

tDAR is a bespoke digital repository designed to provide deposit, access, and preservation services for archaeological data (https://tdar-arch.atlassian.net/wiki/spaces/DEV/overview). During the initial design stages, tDAR was conceptualized not as a digital repository, but as a tool for data integration, that is, a system that allows for the reconciliation or synthesis of datasets with different structures and coding schemes so that they can be analyzed together. It then evolved from a specific research tool into a digital repository that provided more general archival and access services. tDAR was then developed by Digital Antiquity Staff as an open-source software platform, written primarily in Java and JavaScript and licensed under an “Apache 2” open source license, to be a comprehensive archive specifically for resources related to archaeological investigations.

The tDAR application has three main infrastructure components: the PostGres database(s), the filestore, and SOLR. These components were designed by the DA programming staff to provide the primary infrastructure for tDAR. SOLR provides both a search interface to tDAR’s primary entities, and also a caching layer (effectively acting as a document store). The PostGres database(s) function as the working data-layer for tDAR data. The filestore is aimed to be the archival representation of tDAR data. To knit these infrastructure components together, we use Spring to combine components together, and Hibernate as an Object Relational Management (ORM) tool. On top of this, we use Struts 2 to provide the MVC framework and thus manage communications with the browser. tDAR’s frontend uses “Freemarker” which provides a templating language around generating HTML for the browser. Bootstrap 4 is used for the visual appearance management, and VueJS and JQuery for scripting the user interface.

For long-term sustainability, Digital Antiquity works closely with its parent organization, Arizona State University, to ensure that the repository infrastructure is both reliable and stable. The ASU Research Computing Center (https://cores.research.asu.edu/research-computing/about) and Knowledge Enterprise Services (https://research.asu.edu/) provide technical infrastructure, while the ASU Libraries provide a test bed for computing platforms and succession planning.

All tDAR open-source software is run on virtual machines running in ASU’s Research Computing Center. The servers use Ubuntu v18.04 as their base operating system. Regarding availability, bandwidth, and connectivity, tDAR benefits greatly from being part of the ASU Research Computing services that these services provide users with 1 GB download speeds. To facilitate larger downloads, the programming staff can zip large files and compress on-the-fly to assist end-users.

tDAR maintains a regular ‘technology watch’ to ensure that due diligence is paid toward technological changes within its designated community, and, where necessary, hardware and software is maintained to ensure continued operations. Through its relationship with the ASU, tDAR has access to a broad spectrum of software and hardware used within its designated community. Where necessary software/hardware is acquired by the repository. A full inventory of hardware and software is maintained by the Digital Antiquity programming staff to ensure clear documentation of the technical infrastructure of the repository.

As a general guide for standards, tDAR’s design follows the ISO OAIS Reference Model (http://www.oais.info/) and portions of the OAIS Model’s Submission Information Package (SIP), Dissemination Information Package (DIP), and Archival Information Package (AIP) (https://guides.archaeologydataservice.ac.uk/g2gp/App_OAIS).

Under these standards, tDAR is set up so that for each record archived and stored, the system generates an XML representation of that record. The system creates subfolders for each file associated with that record, and the system uses naming conventions that include dates and version numbers. The system generates “derivatives” of record submissions and stores these in a separate subfolder. The system generates these derivatives to aid in the dissemination of a record's content to users. To ensure that this process is followed, the system is also set up to only accept certain file types (https://core.tdar.org/contribute). tDAR users may only submit these file type and encoded validators are used for validation (https://tdar-arch.atlassian.net/wiki/spaces/DEV/pages/720950/tDAR+technical+overview#tDARtechnicaloverview-ThetDARFilestore). From a technical perspective, these standards are implemented using Java scripts

tDAR also follows the W3C standards (https://www.w3.org/standards/) for web design (including HTML, CSS, SVG, Ajax), and XML used for any data interchange.

tDAR also has standards for incorporating third-party software. When selecting a third-party software library for tDAR software development, all selected resources must be open-source and have a software license that allows tDAR to 1) use the library and 2) contribute updates/improvements in return; this is in accordance with our APACHE2 license (http://www.apache.org/licenses/LICENSE-2.0). Digital Antiquity programming staff routinely builds "new" versions of tDAR to ensure that we are using the most recent versions of all libraries (https://tdar-arch.atlassian.net/wiki/spaces/DEV/pages/720950/tDAR+technical+overview).

Digital Antiquity does not currently have a plan in place for future infrastructure development. At present, infrastructure development is informed by the Board of Directors, which meet quarterly to address tDAR operations. The Board has a general policy of monitoring and reviewing the infrastructure in place, and as needs arise, the programming staff makes appropriate changes based on either the Director's or Board of Directors' recommendation(s) or based on service secured by a client. tDAR does employ a JIRA ticketing system that allows Digital Antiquity staff and public users alike to submit questions, bug reports, and feature requests which are reviewed by the repository development team and tDAR archaeological staff. New features and changes are designed and implemented based on this team's recommendations which features, how they should be implemented, and how they are prioritized.

In the future, Digital Antiquity is considering a platform change (i.e., Fedora or Islandora) and is involved with the ASU Libraries Development Team as they migrate their repository. Upon successful completion, DA will evaluate if that platform represents an upgrade over the current cyberinfrastructure. More information about the current development methodology is stored in the tDAR Confluence Wiki (https://tdar-arch.atlassian.net/wiki/spaces/DEV/pages/720938/technical+infrastructure+notes).

All software used and developed at DA is inventoried with system documentation; this includes third-party software deployed. For software developed in-house, Git software is used for version control and change tracking. The source code for this software is stored offsite using two redundant source code repositories: BitBucket (https://bitbucket.org/tdar/) and Github (https://github.com/digital-antiquity/tdar, https://github.com/digital-antiquity).

Third-party software and utilities used to deploy tDAR are inventoried in multiple ways. Primarily, 1Password (https://1password.com/) a security platform, houses and records sensitive data related to services and utilities. This information includes names and contact information for the third-party services, such as domain hosting accounts, storage and archival providers, as well as credentials to sign onto services (https://tdar-arch.atlassian.net/wiki/spaces/DEV/pages/720938/technical+infrastructure+notes). Third-party Java libraries are documented using the Apache Maven dependency management system. When feasible, Digital Antiquity uses community-supported, open-source software. The following core technologies, all of which are open-source community-driven projects, are used extensively 1) Apache Struts 2, 2) Apache SOLR, 3) Apache Maven, 4) JBoss Spring, 5) PostgreSQL, and 6) Git.

Digital Antiquity does have an informal disaster plan in the event of outage; it also has a formal business continuity plan if issues arise with tDAR operations (see attached MOU). These plans ensure that should an issue arise such as cyberinfrastructure failure, or if Digital Antiquity is no longer financially solvent, that tDAR resource owner files will continue to be digitally archived.

Regarding disaster mitigation, which includes catastrophic computing failures, the following steps will be followed.
1. Identify a new, stable hosting service
2. Retrieve most recent copy of source code from Github or Bitbucket
3. Retrieve files and metadata from Amazon Glacier (deep long-term storage); this must be a formal request. Staff would retrieve quarterly backups and daily differential backups, which are also stored on Amazon Glacier.

Reviewer Entry

Reviewer 1

Comments:

Reviewer 2

Comments:

16. Security

R16. The technical infrastructure of the repository provides for protection of the facility and its data, products, services, and users.

Compliance Level:

4 – The guideline has been fully implemented in the repository

Reviewer Entry

Reviewer 1

Comments:
4 – The guideline has been fully implemented in the repository

Reviewer 2

Comments:
4 – The guideline has been fully implemented in the repository
Accept

Response:

Digital Antiquity strives to protect and preserve the archaeological and cultural heritage data and information that is deposited in tDAR. We focus on preserving, curating, and maintaining these data. We accept this responsibility as one of our primary missions (www.digitalantiquity.org/wp-uploads/2017/05/20170503-DA-Security-Summary-Final.pdf). Because of the sensitivity of archaeological information (e.g., culturally sensitive information and geographic information of sites), Digital Antiquity takes security extremely seriously, and incorporates multiple layers of security into our workflows and protocols.

DA employs a Lead Digital Software Engineer (DSE) to lead our security efforts. The DSE is responsible for assessing security risks and dealing with/trouble shooting security issues as they arise. The DSE uses two main risk analysis tools:

1. Atlassian Bamboo (AB; https://www.atlassian.com/software/bamboo) Atlassian Bamboo is a Continuous Integration (CI) tool. Whenever a change to tDAR's source code is made, Bamboo is responsible for checking out a copy of the latest source code, building a temporary, private instance of tDAR based on that source code, and then running a battery of tests against this version of tDAR. These tests verify that tDAR is still in working order, including testing the security features within tDAR. For instance, ensuring that unauthorized data cannot be accessed or inserted into our database, ensuring that confidential and embargoed resources cannot be accessed by unauthorized users. While rare, occasionally a change to one aspect of the source code can introduce a vulnerability in such a way that causes one or more of these automated tests to fail. Bamboo notifies the developers when such test failures occur.
2. GitHub Dependabot (https://docs.github.com/en/code-security/supply-chain-security/about-dependabot-security-updates): Dependabot security updates allows tDAR developers to fix vulnerable dependencies in the repository. When enabled, a Dependabot alert is raised for a vulnerable dependency in the dependency graph of the repository; resulting in Dependabot automatically attempting to "fix" the vulnerability. These alerts occur when the Dependabot utility finds a library referenced in the code that also appears in their vulnerability database. Dependabot will upload suggest security fixes that can be integrated into the tDAR production server.

In terms of physical protection, Digital Antiquity's offices are located in West Hall on the main campus of Arizona State University in Tempe. Access to the offices during business hours is controlled by Digital Antiquity staff. The office area is locked when staff are not present. Computers within the offices are password protected. Access to the data is limited to designated Digital Antiquity data curation staff and management during their time of employment. Access is provided via a secure connection.

Regarding security cyberinfrastructure, Digital Antiquity employs multiple strategies to assure the security of digital files stored in the repository. tDAR uses 256-bit TLS 1.1 encryption throughout the website and application to secure information. After a user registers and logs in to tDAR, all actions occur over a secure channel (e.g. uploading files, making purchases, viewing resources, etc.). The only actions non-registered users can perform are search and view basic metadata for resources and collections in tDAR, non-registered users cannot access data files.

Files are stored on servers at ASU's data center. A suite of tests is run every time Digital Antiquity makes any modification to tDAR's source code. In addition to the testing by Digital Antiquity technical staff for each release, ASU's data center and Digital Antiquity run audits using a suite of common intrusion testing tools to identify potential vulnerabilities.

Files stored in tDAR are protected by multiple, redundant security measures. Physical access to the data center where tDAR's files are stored is restricted and monitored. Data center staff do not have sign-on permissions to tDAR or the virtual machines that run it. In addition, a firewall has been constructed to prevent tDAR's database and backend file store from communicating with any host other than the tDAR webserver.

To protect files from catastrophic loss, Digital Antiquity maintains two backup procedures for tDAR's data. Both procedures employ strong encryption. One set of backups (updated biweekly) are kept in the Phoenix area in a secure storage area. The second set of backups are maintained in the Virginia data center to utilize Amazon's Glacier storage service.

Security and Access Control for Confidential Files

Contributors to tDAR can control and limit access to files they place in the repository. The metadata about those files are always public, which means that anyone can learn about the existence of the resource. Public metadata in tDAR records, such as title and description, but not exact site location or files, are exposed to search engines (e.g., Google, Bing, etc.) for indexing. Digital files in tDAR can be marked as public, confidential, or embargoed. When a digital file is marked "public" anyone who is a registered tDAR user and logged in may download the file. A file marked "confidential" will be inaccessible to users who have not been explicitly granted access to that file by the individual who has this authority. Records in tDAR that have attached confidential or embargoed files provide a link that can be used to request access. Each request generates an email to the record owner. The record owner may decide to provide or not provide access to this request. Digital Antiquity facilitates communication between the record owner and the individual requesting access, but does not grant or deny such requests. Embargoed files are treated as confidential files for a user-designated period, after which the file becomes publicly accessible. Contributors may change the access designation at any time.

The metadata describing tDAR records allow contributors to designate UTM coordinates or specific site locations on a map. If these spatial designations are smaller than one square mile, the tDAR software will obfuscate the spatial data when displaying this information to users who have not logged in or have not been explicitly granted access. When obfuscated, spatial designations will be randomized and will display an area greater than one square mile.

Many clients choose to provide publicly appropriate redacted versions of digital files to upload to tDAR along with full confidential versions containing sensitive information. Digital Antiquity provides redaction services for clients who wish to take advantage of this option. Using professional redaction tools, data curators permanently remove confidential or sensitive information (e.g. archaeological site locations or other information as designated by the client) from a copy of the complete file. This service produces an edited version of the report that is appropriate for access by registered tDAR users.

For an outside evaluation of tDAR's security please review a 2014 report compiled by Sara Rivers Cofield, Curator of the Maryland Archaeological Conservation Laboratory (MAC Lab), who received a Department of Defense Legacy Grant to evaluate tDAR as a repository for digital portions of collections held at the MAC Lab for Defense agencies. Section 5.2 of the report (pp 54-56) addresses the questions related to data security using tDAR (http://core.tdar.org/document/393996/evaluating-a-cooperative-approach-to-the-management-of-digital-archaeological-records).

Reviewer Entry

Reviewer 1

Comments:

Reviewer 2

Comments:

APPLICANT FEEDBACK

Comments/feedback

These Requirements are not seen as final, and we value your input to improve the CoreTrustSeal certification procedure. Any comments on the quality of the Requirements, their relevance to your organization, or any other contribution, will be considered as part of future iterations.

Response:

Reviewer Entry

Reviewer 1

Comments:
The applicant has made the amendments requested.

Reviewer 2

Comments: