The Digital Dilemma

Preservation and the Digital Archaeological Record

Mary Clarke

The assumption that all research ought to contribute to both present and future knowledge is prevalent within various academic disciplines. To ensure that these contributions to the academic commons are in fact viable for the future, scholars have turned to digital formats and methodologies, which are considered to be easier to share, replicate, and store than material formats. However, digital formats are not always more secure. In the fields of archaeology and cultural heritage management, for instance, all-digital workflows that produce intricate networks of associated data containing everything from field notes to virtual worlds are at risk of loss if not properly cared for. To ensure that these digital data formats and the information they contain are accessible and able to function well into the future, scholars will need to navigate the process of digital data preservation in addition to the emerging world of open access academia.

ABSTRACT

The long-term care of collected and created data is an ethical obligation in the fields of archaeology and cultural heritage management. With the growing application of digital methodologies in these fields and the complexity of the resulting data, this task has become complicated. Digital data preservation firms have emerged since this methodological shift, but their policies—championing the democratization of academic data—may conflict with the legal obligations dictated by the countries where data originate. Scholars thus face an inevitable choice between two obligations, one ethical and one legal. While the amount of digital data grows and the options for preservation remain fundamentally misaligned with research norms and project workflows, the digital dilemma places the integrity of data at risk of loss. This article addresses this dilemma by evaluating the existing data publication, archiving, and preservation repositories and considering how, as solutions to the digital dilemma, they can be integrated into multiple workflows. I also propose new directions for archaeological associations, suggesting that they should establish a means of evaluation and approval for third-party preservation firms managing the future of academic research prior to their inevitable ubiquity.

El cuidado a largo plazo de los datos recogidos y creados es una obligación ética en el campo de la arqueología y la gestión del patrimonio cultural. Con la creciente aplicación de metodologías digitales en estos campos y la complejidad de los datos resultantes, esta tarea se ha complicado. Empresas de conservación de datos digitales han surgido desde este cambio metodológico, pero sus políticas—que defienden la democratización de la académica de datos pueden entrar en conflicto con las obligaciones legales impuestas por los países donde se originan los datos. Los estudiosos de este modo se enfrentan a la elección inevitable entre una obligación y otra, una ética y un legal. Mientras que la cantidad de datos digitales crece y las soluciones para su conservación se mantienen fundamentalmente desalineada con las normas de investigación y los flujos de trabajo del proyecto, el dilema digitales coloca la integridad de los datos en riesgo de pérdida. Este artículo aborda las soluciones a este dilema mediante la evaluación de los repositorios existentes para el archivado de datos, la publicación de datos, y la preservación de datos y la forma en que pueden integrarse en múltiples flujos de trabajo. También propongo nuevas direcciones para las asociaciones arqueológicas en que deberían establecer un medio de evaluación y aprobación de las empresas de conservación de terceros que gestionan el futuro de la investigación académica antes de su inevitable ubicuidad.

Advances in Archaeological Practice 3(4), 2015, pp. 313–330
Copyright 2015© The Society for American Archaeology
DOI: 10.7183/2326-3768.3.4.313

https://doi.org/10.7183/2326-3768.3.4.313 Published online by Cambridge University Press

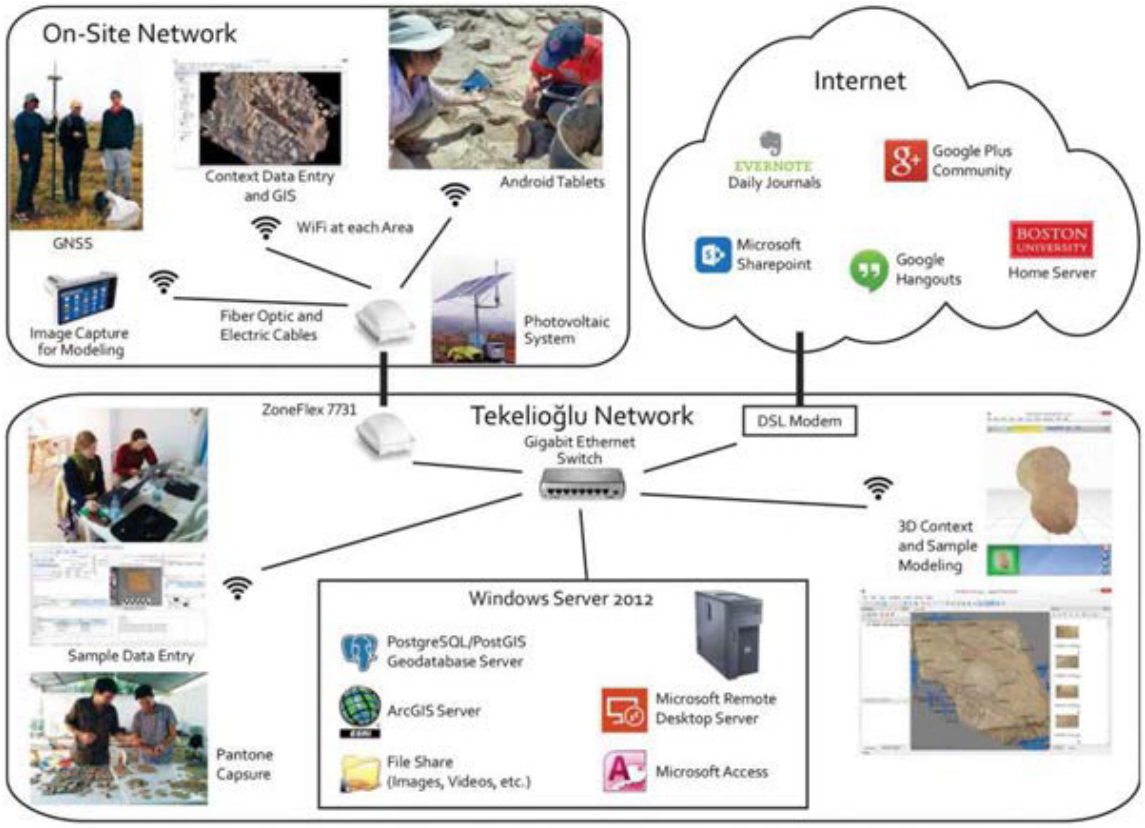


Figure 1. Schematic diagram illustrating the Kaymakçi Archaeological Project's wireless recording and data entry workflow (Roosevelt et al. 2015:Figure 3)

The scholarly future created by all-digital methodologies is one in which preservation firms, not individual academics, become the stewards of scholarship in its digital form. Scholars now need to make the choice between one firm and another while also navigating project-specific legal responsibilities that may conflict with the terms of use and access agreements of the available preservation options. In order to move forward, it is essential that the information needed to navigate these shifts in archaeological practice, and academia more broadly, is located in one place. Accordingly, this article, which frames the issue of digital data preservation within the ethical and legal culture of archaeology and cultural heritage management, reviews data preservation and publication firms in terms of their foci, flexibility, and accepted formats, highlighting how workflows can easily be modified to account for the types of data outputs created by scholars today.

THE DIGITAL DILEMMA

Digital data, while easy to store, copy, and share, need constant management (Altman et al. 2009; Berman 2008; Conway 2010; Conway et al. 2011; Dürr et al. 2008; Gunia and Sandusky 2010; Hockx-Yu and Knight 2008; Jeffrey 2012; Rieger 2008). Data management addresses issues such as file corruption, file and media obsolescence, and inadequate metadata (Jeffrey 2012:556); data migration when software or data become unreadable (Cultural Heritage Partners 2012; Jeffrey 2012); and data curation, which supports the work of researchers, present and future, who should be able to search, cite, and reuse[1] earlier data (Alexander 2013; Faniel et al. 2013; Hanson et al. 2011; Jeffrey 2012; Kansa and Kansa 2013; Kansa and Kansa 2014; King 2011; Molloy 2011; Porter 2013). Dumping or "backing up" all project data into an archive or on a storage device is not an adequate data preservation strategy (Baker and Yarmey 2009). Instead, digital data preservation is the process of making primary data reusable in the sense that it both continues to function as operating systems evolve and also contains sufficient metadata detailing its creation and history (Archaeology Data Services and Digital Antiquity 2011; Faniel et al. 2013; Kansa 2012; Kansa and Kansa 2013; OpenContext 2015; Porter 2013; Richards and Winters 2015). The world of open access academia and data publishing is shaping the future of digital data curation and reuse, and data preservation is part of this movement (Baker and Yarmey 2009; Conway et al. 2011; Cultural Heritage Partners 2012; Dürr 2008; Hall 2013). While the amount of digital data grows and the solutions for their preservation remain fundamentally misaligned with existing workflows and proprietary claims, both individually and internationally, the digital archaeological record is at risk of loss (Petrovic et al. 2011). If scholars wish to ensure the long-term viability of their digital data, they must navigate the potential conflicts that arise from these advances. This is the digital dilemma faced by scholars advocating for digital and paperless workflows today: preserve and publish your data in an open access world, or watch it perish.

https://doi.org/10.7183/2326-3768.3.4.313 Published online by Cambridge University Press



Figure 2. The process of paperless field illustration as designed by Forte (2014) and the Çatahöyük Project (images courtesy of the 3D-Digging at Çatahöyük Project).

DIGITAL ARCHAEOLOGY

Proponents of digital methodologies in archaeology frequently cite terms such as "archive," "preservation," and "conservation" of cultural heritage, with an added focus on data reuse by both the public and future researchers (Barton 2009; Beale 2012; Beaubien et al. 2007; Bruno et al. 2010; CyArk 2003, 2015; De Reu et al. 2014; Eve 2012; Forte 2014; Green et al. 2014; Guidi et al. 2014; Johnson 2012; Meyer et al. 2007; Remondino et al. 2009; Roosevelt et al. 2015; Sanders 2014; Shaw 2012). This idea of permanence or an enduring record through the adoption of digital methodologies is pervasive, yet future plans for "born digital" data are infrequently cited. While some data do not appear to pose an immediate challenge for future preservation, other methodologies, and frequently those that make the largest claims for their archival capabilities, appear to be the most at risk of future obsolescence (Jeffrey 2012).

Field Notes, Illustrations, and Reports

While many archaeological projects still use standard materials for in-field documentation, advocates for fully digital, "paperless" workflows are pushing the discipline in a new direction (DeRue et al. 2014; Forte 2014; Roosevelt et al. 2015). For example, on their project at the site of Kaymakçi, Roosevelt et al. (2015) implemented an entirely paperless workflow in which project archaeologists take notes exclusively on tablets that sync to and are stored on a central cloud in real time (Figure 1). This workflow allows the project directors to check in on excavation units across the survey area and to view notes as they are being written. Similarly, at nearby Çatalhöyük, Maurizio Forte (2014) uses a workflow that bypasses the need for hand-drawn illustrations of excavation units (Figure 2). By tracing georeferenced photographs of test pits directly onto a tablet and subsequently referring back to the units for any needed corrections, Forte has reduced documentation time and eliminated costs for traditional materials (Figure 3) (Forte 2014:15). Printed monographs and final report publications have also become a rarity within the field. Instead, projects are choosing to publish documents digitally on open access websites (Çatalhöyük Research Project 2015; Leon Levy Expedition to Ashkelon 2015).

The abovementioned all-digital workflows change the format of the archaeological record: raw data and analyses, such as field notes, photographs, illustrations, and subsequent reports, are now "born digital" documents in need of long-term care. While these are arguably the types of data least at risk of obsolescence in that they can be printed and rescanned, they are at risk of loss if file associations and metadata are not maintained. As the amount of data generated from an entire archaeological project can be huge, and the associations between files may be understood by only a few project members, the future reuse of archaeological data is at risk if it is not properly organized or curated, which is often the case when data is merely dumped into a repository (Baker and Yarmey 2009). Furthermore, data reuse requires detailed metadata (Tables 1 and 2) so that future scholars are able to determine any potential biases in initial collection as well as the quality of the resulting analyses. If scholars want to prevent information loss and the obsolescence of their digital data, they need to advance their workflows so that they incorporate metadata authorship as well as preservation.

Three-Dimensional Imaging

Considered the most accurate and enduring method of documentation among many archaeologists today, three-dimensional (3D) imaging technologies are able to create detailed replications of archaeological materials and world heritage monuments (Armstrong 2006; Barton 2009; Beaubien 2007; Bruno et al. 2009; CyArk 2003, 2015; DeReu et al. 2014; Forte 2014; Green et al. 2014; Guidi et al. 2014; Remondino et al. 2009; Roosevelt et al. 2015; Sanders 2014; Shaw 2012). For example, an organization named CyArk (2003, 2015) made the bold commitment to digitize all of the UNESCO World Heritage sites using LiDAR and other advanced scanning technologies (Figures 4 and 5). Their mission states that CyArk was founded "to ensure heritage sites are available to future generations … with the mission of using new technologies to create a free, 3D online library of the world's cultural heritage sites before they are lost to natural

https://doi.org/10.7183/2326-3768.3.4.313 Published online by Cambridge University Press

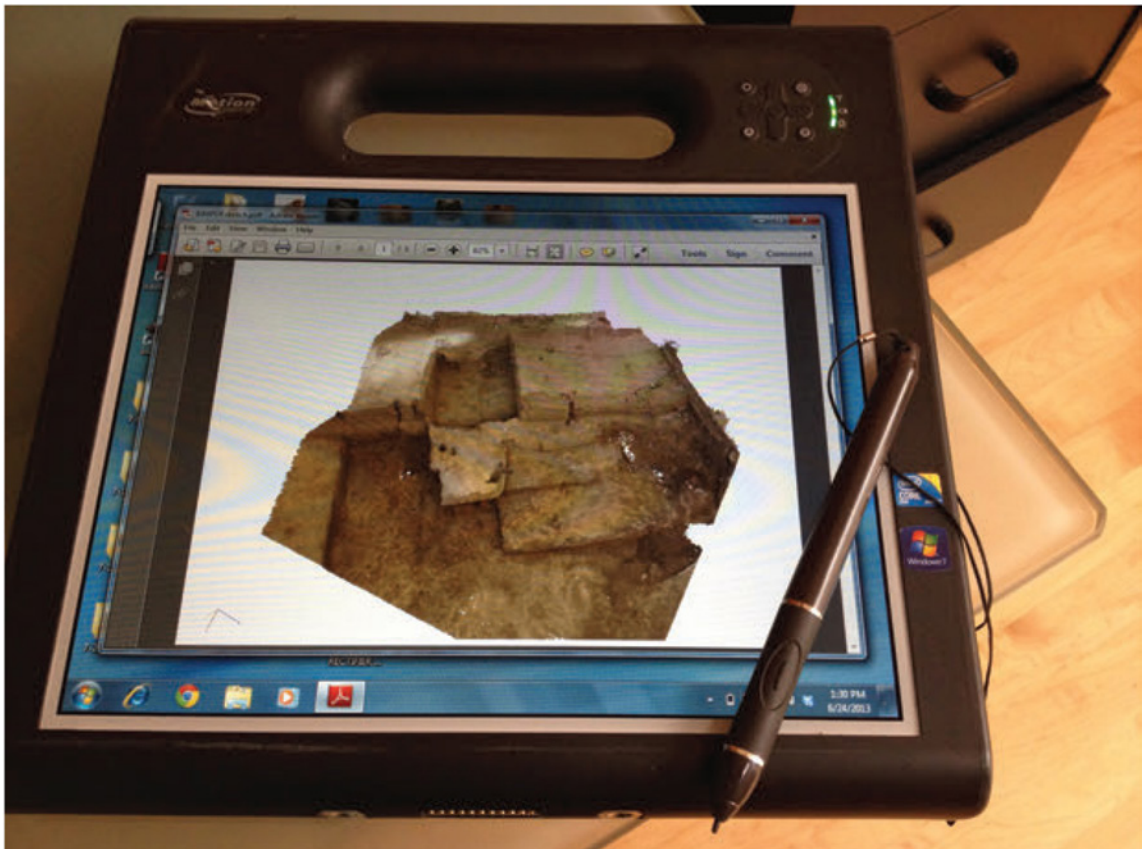


Figure 3. The process of finalized field illustrations as designed by Forte (2014) and the Çatahöyük Project (image courtesy of the 3D-Digging at Çatahöyük Project).

disasters, destroyed by human aggression or ravaged by the passing of time" (CyArk 2003). With the application of scanning technologies, scholars are able to digitally preserve cultural landmarks and their associated knowledge for posterity, a feat possible only with these advanced technologies. However, the future accessibility and long-term preservation of digital cultural heritage monuments, and not just their real-world counterparts, must also be addressed.

Certainly documenting and providing access to these monuments should be a primary concern (Beale 2012), but we also need future plans for both access to and care of the data themselves. While not exactly complex, the datasets associated with image captures for 3D documentation are quite large. Each completed scan is composed of a network of polygons (Figure 4) that knit together to form the shape or geometry of the scanned monument. The mesh itself does not contain image files, but at each vertex, a photograph is taken, creating a secondary texture file. When these two lines of data are joined together—the texture file lying on top of the shape file—the scanned monument can be observed and studied.

Long-term care of these datasets requires that each individual file composing the final replica be maintained in an operable and uncorrupted state (Archaeological Data Services and Digital Antiquity 2011). Files go through several preservation intervention points where they fundamentally transform in format during the process of becoming a final version of a scan or 3D model, known as an asset (Archaeological Data Services and Digital Antiquity 2011). Preserving these datasets requires that "copies of the data should be archived after each significant step in the process" (Archaeological Data Services and Digital Antiquity 2011). Metadata detailing the creation and location of an individual file within the mesh also must always follow a completed asset (Table 2). If metadata are maintained carefully, data managers can go into the shape or image files and repair a single broken link and also migrate data as operating systems or proprietary claims change. Thus, for the preservation of big data, all files generated by 3D-scanning technologies as well as all the permutations resulting from the creation process need to be preserved.

Virtual Reality

The use of virtual reality in the fields of archaeology and cultural heritage management is quite different from the other workflows described in this article. Where paperless excavations and 3D imaging are methodologies used to collect or capture primary data, virtual realities are interpretations of data rather than data themselves (Bawaya 2010; Bonde and Houston 2013; Bruno 2010; Dassault Systems 2013, 2015; Eve 2012; Favro 2013; Gillings 2005; Guidi 2014; Johnson 2012; Mesick 2013; Museum of Fine Arts Boston 2007; Sanders 2014). However, they are interpretations based on archaeological data for the purpose of

https://doi.org/10.7183/2326-3768.3.4.313 Published online by Cambridge University Press

TABLE 1. An Example of the Information Typically Deposited with Collections of Project Data, from the Archaeological Data Service and Digital Antiquity's Metadata Project.

Element	Description
Project Title	The title (and any alternatives) for the dataset.
Description	A brief summary of the main aims and objectives of the research project (or alternative process) from which the data collection arose together with a brief summary description of the content of the dataset.
Subject	Keywords for the subject content of the dataset (qualified using, e.g., the English Heritage NMR Monument Type Thesaurus or the MDA Object Type Thesaurus.
Coverage	This is both spatial and temporal coverage. For spatial coverage it should include the current and contemporary name(s) of the country, region, county, town or village covered by the data collection and, where possible, a standardized reference such as the Getty Thesaurus of Geographic Names should be used. If names or administrative units were different during the time period covered by the data they should be recorded separately. Site coordinates can also be entered as a National grid reference in a number of different ways e.g., as a point (useful to describe a small project area via a central coordinate); as a line (e.g., at least 2 coordinates to represent the linear limits of the site); as a polygon (for a more complex site area, 3 or more coordinates are used to describe the boundaries). If applicable, the full postal code for the site can be included. For temporal coverage, it should include the dates/period covered by the dataset (using existing thesauri where possible such as the RCHME Period List).
Creators	Details of the creator(s), compiler(s), funding agencies, or other bodies or people intellectually responsible for the data collection. Information should include forename, surname, affiliation, address, phone, fax, email, or URL.
Publisher	Details about any organization that has published this data.
Contributors	Other individuals or organizations that have contributed to the resource.
Identifiers	Project or reference numbers used to identify the dataset.
Source	Any important earlier work(s) from which this resource is derived.
Dates	Dates indicating when the dataset was created, when the archaeological project was carried out, processing dates, or computerization dates as appropriate.
Copyright	The name of the copyright holder for the dataset. If the collection was created during work by an employee, the copyright holder will normally be the employer. If the material is covered by a specific copyright (e.g., Crown copyright) please indicate this.
Relations	If the data collection was derived in whole or in part from published or unpublished sources, whether printed or machine-readable, this element should include references to the original material, details of where the sources are held and how they are identified there (e.g., by accession number). If the collection is derived from other sources, include an indication of whether the data represents a complete or partial transcription/copy and the methodology used for its digitization. Also, include full references to any publications about or based upon the data collection.
Language	Indication of which language(s) the dataset is in (e.g., English, French, Spanish).
Resource Type	Whether the dataset is best described as primary data, processed data, an interpretation of data, or a final report.
Format	The format the data is saved in (e.g., WordPerfect 5.1, HTML, AutoCAD).

Note: From Archaeological Data Services and Digital Antiquity's Guidelines for Good Practices.

visualizing an invisible past. The contributions of virtual reality to the field of archaeology, specifically to community or public archaeology, can clearly be seen at the extensive visualization project Giza3D (Figure 6) (Bawaya 2010; Dassault Systems 2013; Shaw 2012). Beginning first as an archive for the digitized copies of 40 or so years of field notes, illustrations, and maps, Giza3D is now an interactive environment built from archaeological data (Museum of Fine Arts Boston 2007).[2] Every excavated element of Giza was reconstructed using project data, and along with every reconstruction of a tomb (or *mastaba*), the user can view the handwritten notes of the excavator, illustrations of stratigraphy, and every other associated file (Figure 7). Furthermore, Giza3D was structured for public access. A user simply logs on to the website and, if in possession of the appropriate software, can begin independent exploration of an animated antiquity.

The consolidation, unification, and visualization of primary archaeological data inspire the application of virtual methodologies, as these methods can subsequently be used for community engagement, education, and tourism, in addition

https://doi.org/10.7183/2326-3768.3.4.313 Published online by Cambridge University Press

TABLE 2. An Example of the Information Typically Deposited with Individual Files within Project collections, from the Archaeological Data Service and Digital Antiquity's Metadata Project.

Element	Description
File name	The name of the file, e.g., report.doc
File format	The file format, e.g., PDF/A or Open Office Document
Software used to create the files	The software used to create the file, e.g., Microsoft Word 2007
Hardware used to create the files	The hardware used to create the file; this is more significant when files are created directly by survey equipment such as laser scanners or GPS devices.
Operating system used to create the files	The operating system under which the file was made, e.g., Windows XP or Mac OS X 10.5.
Date of creation/last file update	When the file was made or updated.
Processing history or lineage	This element should be used to highlight relationships between files and whether a file is a source file or derived from another.

Note: From Archaeological Data Services and Digital Antiquity's Guidelines for Good Practices.

to phenomenological approaches to ancient experiences. (Beale 2012; Meyer et al. 2007). As visualizations of archaeological research, these environments, or virtual worlds, should be published with the same degree of transparency as any written interpretation. These types of data should be archived with the raw data used to inform the interpretation as well as the multiple versions of included assets and their metadata, so that future scholars wishing to cite, reuse, or further contribute to the environments can confidently do so. As alternative archive platforms, these environments need extensive data management structures that curate and preserve all associated data while maintaining user operability on all manners of interfaces (Kintigh et al. 2015). Because these data are big and complex, and the mismanagement of their organizational structures could destabilize their integrity, investing in their long-term maintenance is as important as investing in their production. The Giza Project, for instance, must remain vigilant to upgrade its virtual reality environment for maximum access to users. According to project director Peter Der Manuelian (personal communication, 2015), the Giza Project at Harvard University "must balance the need for open source software (often preferred by granting agencies) with the development of leading-edge innovative features that make the Giza content so compelling." As this project integrates Giza's entire excavation history with its virtual representation, this is not a simple task. However, to neglect its long-term care would essentially undermine the Project's intent: the publication and unification of Giza's complete archaeological record.

The risk of digital data, whether in the form of field notes or virtual worlds, becoming unusable if not suitably cared for is becoming increasingly apparent. Thus, *whether* we should preserve and curate our new digital archaeological record is no longer in question. Instead, the questions we must consider are *who* should be responsible, *when* in our workflows should we begin the process, *what* exactly is worth preserving, and *how* should we address future access and preservation for the long term?



Figure 4. LiDAR image of Mt. Rushmore produced by CyArk (image courtesy of CyArk archives).

https://doi.org/10.7183/2326-3768.3.4.313 Published online by Cambridge University Press



Figure 5. The process of scanning UNESCO World Heritage sites (image courtesy of Cyark Archives).

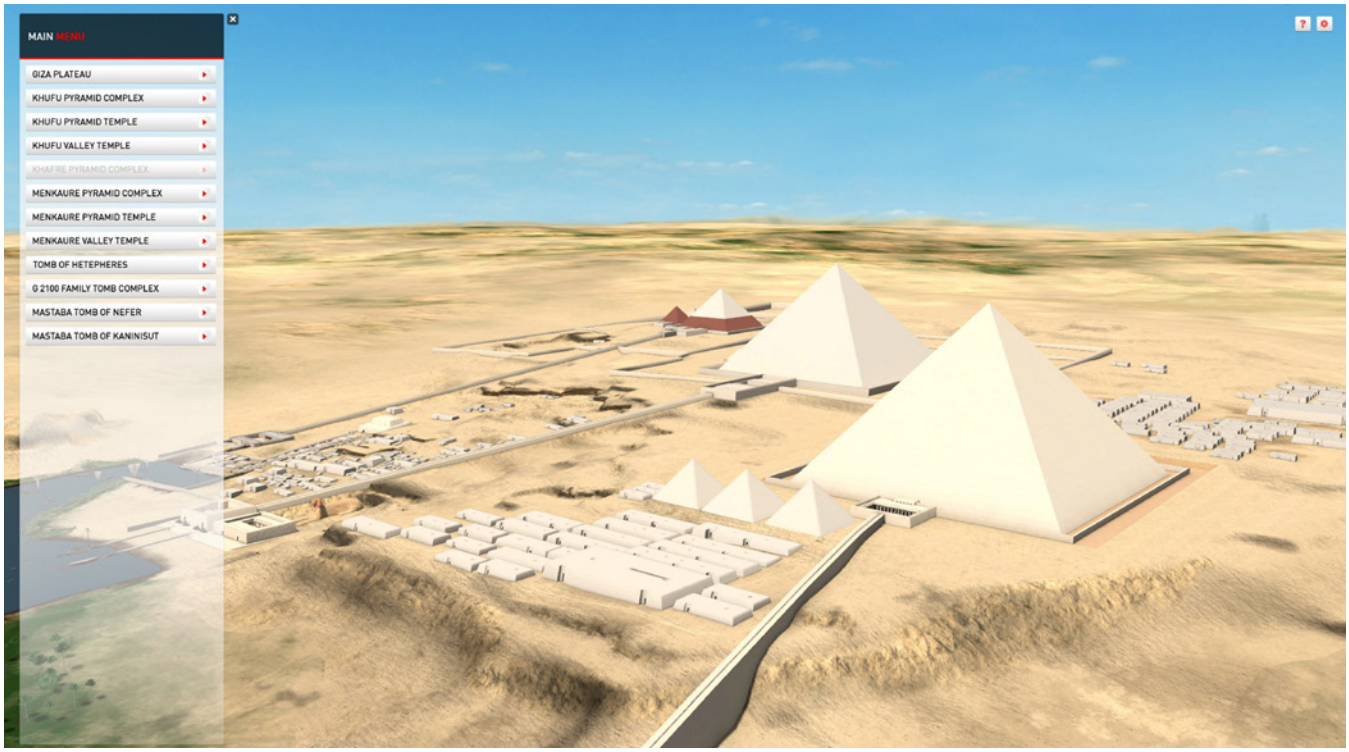


Figure 6. Giza3D (image Courtesy of Dassault Systems).

https://doi.org/10.7183/2326-3768.3.4.313 Published online by Cambridge University Press

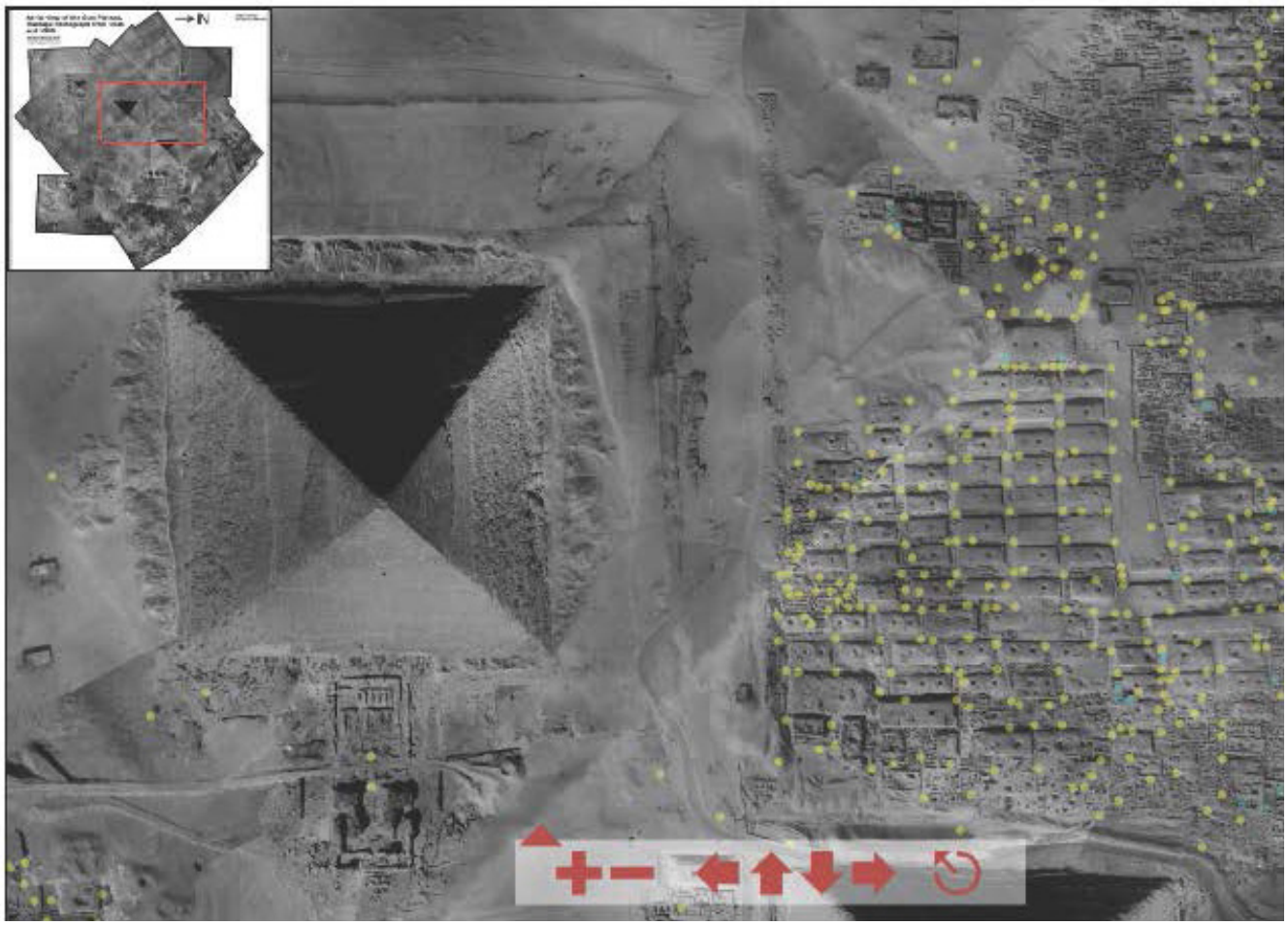


Figure 7. Interactive archive of Giza's excavation history hosted by the Museum of Fine Arts, Boston. Within the image, users scroll over architectural features, at which point they are directed to an inventory of the associated archaeological data. (Image courtesy of Museum of Fine Arts, Boston Giza Project).

THE ETHICAL AND LEGAL ARGUMENT FOR RESPONSIBILITY OR STEWARDSHIP

Responsibility or stewardship of digital data is complicated. As stewards of the archaeological record, archaeologists should also be stewards of the digital archaeological record, but attributive questions of responsibility are problematic and plagued with confusion and potential legal issues. First and foremost in this debate is the question of *who* should pay or be responsible for the long-term preservation of digital data (Alexander 2013; Berman 2008; Hall 2013; Kansa 2012; Kansa and Kansa 2014; Kintigh et al. 2015; Porter 2013; Pratt 2013). Should individual archaeologists pay only for the preservation of their current and future digital data, or should we also be responsible for retroactively preserving all significant datasets created in the past? If long-term digital preservation of archaeological data is left up to individual scholars, chances are it will be neglected as other more pressing concerns, such as funding and tenure, take priority.

As detailed above, the digital data in question are vast; thus, we should also evaluate *which* data warrant the cost of long-term care. On this issue, federal laws are explicit for those working within the United States or receiving federal or national funding (Cultural Heritage Partners 2012; National Science Foundation 2013, 2015). Conversely, associations disseminating codes of ethics and principles, such as the Archaeological Institute of America (AIA 1997), the Register for Professional Archaeologists (RPA 1998), the Society for American Archaeology (SAA 1996), and the Society for Historical Archaeology (SHA 2003), are less clear. While the SAA and the Digital Archaeological Record (SAA and tDAR 2013) have recently addressed the issue of *how* to preserve digital data by advocating for a specific preservation firm, the Digital Archaeological Record (tDAR 2015), most have yet to provide clear solutions associated with the changes in contemporary archaeological practice.

The Archaeological Institute of America

Section 1.4 of the AIA's Code of Professional Standards states, "Archaeologists should anticipate and provide for adequate

https://doi.org/10.7183/2326-3768.3.4.313 Published online by Cambridge University Press

and accessible long-term storage and curatorial facilities for all archaeological materials, records, and archives, including machine-readable data, which require specialized archival care and maintenance" (AIA 1997). The additional clarification of "machine readable data" is unique to the AIA. They appear to be the only association that has made a concerted effort to identify a specific conservation need for specialized data, such as digital data. Furthermore, section 1.6 goes on to attribute responsibility of preservation to the research project and to set a time frame for conservation plans. It asserts that all research projects "should contain specific plans for conservation, preservation, and publication from the very outset, and funds should be secured for such purposes" (AIA 1997).

The Society for American Archaeology

The SAA's (1996) Principles of Archaeological Ethics contain applicable principles. In defining "Stewardship," the first principle states,

> It is the responsibility of all archaeologists to work for the long-term conservation and protection of the archaeological record by practicing and promoting stewardship of the archaeological record. Stewards are both caretakers of and advocates for the archaeological record for the benefit of all people; as they investigate and interpret the record, they should use the specialized knowledge they gain to … support … its long-term preservation [SAA 1996].

As expressed here, archaeologists hold privileged positions as caretakers of the archaeological record and are ethically obligated to support the long-term preservation of what they interpret and record. Principle 6 adds that "documents and materials on which publication and other forms of public reporting are based should be deposited in a suitable place for permanent safekeeping" (SAA 1996). In other words, data used to construct interpretations, including individual excavation photos and digitized illustrations, ought to be preserved for the long-term.

Principle 7, on the other hand, takes a soft approach to data conservation. It states, "Archaeologists should work actively for the preservation of, and long-term access to, archaeological collections, records, and reports" (SAA 1996). With a larger focus on in situ conservation, the SAA has not modified its vernacular for digital methodologies. The SAA should make revisions to its principles to make explicit their take on the ethical obligations of digital data preservation.

The Society for Historical Archaeology and the Register of Professional Archaeologists

The SHA (2003) and the RPA (1998) similarly have vague articles referring to the ethical obligations to care for digital data.[6] The SHA discusses "long-term preservation … of sites and collections … for the benefit of humanity" in its Principle 2, and then Principle 4 asserts that "members of the Society for Historical Archaeology have a duty to collect data accurately … and to see that these materials are appropriately curated for future generations" (SHA 2003). Certainly it can be inferred that these principles apply to the preservation of digital data and heritage, but the lack of direct recognition should be modified. The RPA (1998) is far more problematic. The closest it comes to mentioning stewardship or future plans for excavated material is under section 3.5, where it states, "Specimens and research records resulting from a project must be deposited at an institution with permanent curatorial facilities, unless otherwise required by law" (RPA 1998). In this case, it assigns no responsibility for the material to the excavator or to the project, and passes responsibility off to some unnamed institution and set of laws.

Who Says What About the Long-Term Care of Digital Data?

A clear consensus regarding *who* should take responsibility for data produced through (not extracted during) archaeological excavation cannot be found in current ethical codes and principles. According to the AIA (1997), research projects should be held responsible for the plan of data preservation. The use of "long-term" (AIA 1997; SAA 1996; SHA 2003) suggests that the responsibility of access, preservation, and storage holds for the foreseeable future.

What exactly should be conserved is also ambiguous. The AIA (1997) take a modern stance in that they address the idea that some data, specifically "machine-readable data" require "specialized archival care and maintenance" (AIA 1997). Their definition of data is inclusive, whereas those of the SAA (1996), SHA (2003), and RPA (1998) are composed of broad terms such as collections, material, records, reports, and specimens. The SAA (1996) does place interpretations of data into the larger idea of the archaeological record. It should follow that visual reconstructions are accounted for under this principle. *When* data should be conserved is another issue. Only the AIA (1997) explicitly states that both the plans and funds for long-term care should be defined prior to excavation. The RPA (1998) policy is the weakest, but the SAA (1996) and the SHA (2003) are equally noncommittal in regard to the practicality of data preservation.

No organization directly addresses *how* archaeological projects should approach the long-term care of their digital data within their current ethical codes and principles. The SAA (1996) implies that there should be some form of accessibility to this data; the AIA (1997) suggests specialized care; and the SHA (2003) and RPA (1998) both appear to see outside institutions as potential aides in the future care of archaeological data. They do not explicitly define a project's ethical responsibilities for the long-term future of digital datasets. However, the SAA and tDAR (2013) recently published the following separate statement:

> As digital data become an increasingly important component of archaeological records and documentation, and an integral part of every phase of archaeological research, SAA affirms the ethical responsibility of archaeologists, public agencies, and other organizations funding, authorizing, managing, or conducting archaeological work to ensure long-term preservation of and access to the digital records of their investigations by employing, mandating, or budgeting for use of digital repositories that provide appropriate online access and high standards for long-term preservation [SAA and tDAR 2013].

https://doi.org/10.7183/2326-3768.3.4.313 Published online by Cambridge University Press

This statement by the SAA and tDAR (2013) should be incorporated into the SAA's archaeological principles and should serve as a model for other archaeological associations. While the AIA (1997) takes the best approach from within its published mandates, all of the reviewed associations, as well as others not reviewed here (American Anthropological Association 2012; American Cultural Resources Association 1996; SAA 2009), should modify their existing codes of ethics and principles so that they reflect changes in archaeological practice and directly address the preservation of the digital archaeological record.

Federal Law and the Digital Archaeological Record

In 2012, Arizona State University commissioned Cultural Heritage Partners, PLLC, to look over the existing federal laws pertaining to archaeological and cultural heritage data. The goal of this review was to assess the legal responsibility of archaeologists to preserve and provide long-term access to digital data (Cultural Heritage Partners 2012). Citing the National Historic Preservation Act (NHPA) and the Archaeological Resources Protection Act (ARPA), they established "that the NHPA and ARPA require that archaeological data be maintained permanently in appropriate data bases, made available to potential users, and deposited in an institution with adequate long-term curatorial capabilities" (Advisory Council on Historic Preservation 2014; Cultural Heritage Partners 2012:10; National Park Service 2006 [1979]). The term "adequate" is further explored throughout the text in that institutions must have strategies for data migration (Cultural Heritage Partners 2012:7–9; see below for information on data migration). While the federal laws explored in this document specifically pertain to research conducted within the United States, research funded by national or federal granting bodies is also required to comply, making long-term access to and preservation of digital data a legal as well as an ethical responsibility (Advisory Council on Historic Preservation 2014; Cultural Heritage Partners 2012:10; National Park Service 2006[1979]; National Science Foundation 2013, 2015).

Although most scholars have yet to incorporate digital data preservation into existing workflows, recent changes in academia and federal funding have made it so that there is no longer a choice, and scholars are now legally responsible for providing the access to and preservation of research data. The issue now is *how* should one determine which digital preservation or publication service is worth trusting? In the statement by the SAA and tDAR (2013), the SAA advocates for "repositories that provide appropriate online access and high standards for long-term preservation," but it leaves the interpretation of "appropriate" access and "high standards" up to the individual. While the SAA supports the use of Digital Antiquity's preservation service, tDAR, it does not provide a means for evaluating the many other options (Appendix; SAA and tDAR 2013). The review conducted by Arizona State University and Cultural Heritage Partners, PLLC, illuminated what is federally required of an archive (Cultural Heritage Partners 2012). Other governing bodies, such as the Data Seal of Approval (DSA), bestow distinction upon existing options (Data Archiving and Networked Systems 2005; Institutional Organization for Standardization 2011; National Archives Records Administration 1998). Scholars, however, need to be able to determine on their own which repositories observe high standards and provide the appropriate level of access required by ethical and legal mandates. These semantic issues are important in the context of the growing market for digital data preservation services, and archaeologists would be better equipped to navigate the burgeoning open access world if archaeological associations had an established means of evaluating the available options.

SOLUTIONS TO THE DIGITAL DILEMMA

The digital dilemma faced by scholars today is that they are ethically and legally required to preserve and provide access to their research within the democratic world of open access academia or else watch their research decay. Both individual and project researchers are often wary about digital data preservation because of the perception that it involves a forfeit of rights. Yet this does not have to be the case. Solutions to the digital dilemma are multi-scalar and flexible, and there is likely one that fits or responds to a project's specific legal obligations or an individual's level of comfort (see the Appendix, where current solutions are organized by their foci, licenses, accepted formats, costs, access, future sustainability, guides to best or good practices, and certification or reference models). In addressing this dilemma, the differences between data archiving, data publishing, and data preservation need to be reviewed in relationship to the foci, structure, costs, and sustainability of the available short- and long-term solutions.

Private Archives

Improving a project's workflow, so that it appropriately prepares for the future reuse and long-term care of digital data, can begin with a private archive (e.g., Dropbox, iCloud, etc.) or storage system (e.g., internal or external hard drive). As stated earlier, digital data lose their value if separated from their metadata, thus archaeologists can prevent this type of loss by integrating the composition of metadata into existing archival or storage workflows (Gunia and Sandusky 2010). As seen in Tables 1 and 2, metadata extensively documents the creation and history of a file so that it can be preserved and reused by future scholars (Archaeological Data Services and Digital Antiquity 2011; Dublin Core Metadata Initiative 2015; Library of Congress 2015). This small step eliminates unnecessary redundancy and helps scholars ensure that the information needed for future citations, cross-archive searches (interoperability), and data reuse is maintained. When preparing big or complex digital data such as a 3D model or scan used within virtual worlds, archaeologists should add a description of an asset's worked history to its associated metadata (see Tables 1 and 2 for a description of what information is typically included in metadata), as the ontology or worked history of a file may help diagnose issues with future operability. While the private archive or storage system does not satisfy the ethical or legal responsibilities of archaeologists and cultural heritage managers and cannot be considered a preservation strategy, it does help scholars prepare their data for the next step, which is preserving and/or publishing their digital data.

Data Publishing

Data publishing, an alternative to private-access archives, is a polarized topic within academia. The movement advocates for

https://doi.org/10.7183/2326-3768.3.4.313 Published online by Cambridge University Press

open access, no-rights-reserved sharing of primary data under the Creative Commons (CC) Copyright laws (Creative Commons 2015).[3] Proponents of data publishing see it as an opportunity to change the way academics access, publish, and reuse primary data, not to mention collaborate with other scholars and provide knowledge to a global community (Alexander 2013; Baker and Yarmey 2009; Beale 2012; Borgman 2008; Conway 2011; Faniel et al. 2013; Hall 2013; Hanson et al. 2011; Kansa 2012; Kansa and Kansa 2013; Kansa and Kansa 2014; King 2011; Kintigh et al. 2015; Molloy 2011; Porter 2013; Pratt 2013; Richards and Winters 2015; Zuiderwijk et al. 2012). However, critics of data publishing, and of open academia in general, cite its potential to negatively impact the peer review structure of established scholarship, flood the field with unregulated content with little to no quality control, and diminish a scholar's rights to their own data. Formed in response to these critiques, an organization called OpenContext (2015) has set a new standard for data publishing (Kansa 2012; Kansa and Kansa 2013; Kansa and Kansa 2014; OpenContext 2015). At OpenContext, scholars must first submit their raw data for peer review, where it is judged based on its "methodological soundness and data quality, quality of documentation, and suitability for wider reuse" (OpenContext 2015). If accepted, scholars work with OpenContext to refine their data and author metadata and to determine links between associated files, all in an attempt to ensure ease of reuse and the highest levels of interoperability. Furthermore, citation standards are organized by Digital Object Identifiers (DOIs) so that every individual file has its own stable link, but published collections of data also become citable, peer-reviewed works and can be organized with other published materials during tenure review cases. While the primary aims of OpenContext's data publishing are data reuse, transparent scholarship, and academic collaboration (Creative Commons 2015), they do allow data creators to determine the copyright and licensing agreements on individual items published within their repository (OpenContext 2015).

Other repositories such as the Dataverse Project (2015), the Pleiades Project (2014), SketchFab (2015), Visual Past (2015), and Zenodo (2015) offer similar conditions for digital data publishing, both for small data (e.g., text, image, spreadsheets, etc.) and big or complex data (e.g., geospatial, virtual models, etc.), but they often do so without the peer review structure or without ensuring the long-term viability of the published data (see the Appendix). Likewise, OpenContext, while backed by the California Digital Library, does not presently preserve deposited data. It does, however, make it interoperable with preservation services such as Digital Antiquity's the Digital Archaeological Record (tDAR) and the UK-based Archaeology Data Services (ADS) (OpenContext 2015).[4] If a repository does not preserve your data, but makes that data interoperable through migratory languages (e.g., XML, HTML, ArchaeoML, etc.), you would be able to send your data in for repairs, so to speak, or make that data available in one repository while preserving it in another. Thus, services such as Dataverse (2015) and OpenContext (2015), which publish digital data and make that data interoperable, satisfy the ethical and legal responsibilities of archaeologists and can be considered proper solutions to the digital dilemma when incorporated into project workflows.

Data Preservation

Digital data preservation fits easily into existing project workflows. Preservation services that offer flexible copyright and terms of access agreements, such as Archaeological Data Service (2014a), ADSeasy (Archaeological Data Service 2014b), Figshare (2015), and tDAR (2015), can replace the private archive or storage device while offering proprietary security (see the Licenses and Access sections of the Appendix).[5] In the same way that projects control access to and use of data within Dropbox, iCloud, or hard drive archives, they can control access to and use of digital data curated and preserved in these repositories. There need not be any difference in a project's workflow other than the location of its digital data. While there are publication and preservation services that require individuals and projects to publish data under Creative Commons and no-rights-reserved licenses (Data Archiving and Networking Services 2005, 2014; Dataverse Project 2015; Digital Repository of Ireland 2015; Pleiades Project 2014; Visual Past 2015), and most flexible repositories advocate for open access data publishing and promote the democratization of scholarship (Archaeological Data Service 2014a; ADSeasy [Archaeological Data Service 2014b]; Figshare 2015; tDAR 2015), individual and project researchers wishing to preserve digital data outside of the commons are able to do so. Furthermore, when ready to make data openly available for access, citation, collaboration, and reuse, researchers simply need to change these agreements, which can be done on individual files as well as project collections at any time. While the costs, copyrights, terms of access, and file formats currently being accepted by repositories should give researchers an idea of what option is best for them, so should the availability of guides to good or best practices (Archaeological Data Services and Digital Antiquity 2011), certifications (Data Archiving and Networking Services 2005; International Organization for Standardization 2011; National Archives Records Administration 1998), and reference models (Consultative Committee for Space Data Systems 2002, 2012; CLOCKSS Archive 2015; International Council of Museums 2014; L–P: Archaeology 2015).[6] Such resources often speak to the overall sustainability and future of the archive itself.

DISCUSSION

Integrating preservation strategies into existing project workflows need not be complicated. Individual or project researchers can either prepare their data for future reuse during their existing storage practices or replace their private archive with a repository that actively preserves data under flexible copyright laws. Data publication and preservation services are the solution to the digital dilemma, as they present the only opportunity for access to and reuse of the archaeological record for perpetuity. Based on a review of federal law and the existing principles and codes of ethics, we are legally and ethically responsible for the long-term access and preservation of digital data, but the associations promoting ethical codes need to make this more explicit.

While the SAA has made a significant advance in its joint statement with tDAR (SAA and tDAR 2013), all associations need to amend their codes and principles to include a direct definition of the ethical obligation to preserve digital data as well as devise a means of evaluating the existing solutions. The difference between archiving, publishing, and preserving data collected and created during archaeological research and cultural heritage management could be seen as a semantic one, but, in reality, the difference is a matter of taking a passive or active role in the long-term preservation of the archaeological record.

https://doi.org/10.7183/2326-3768.3.4.313 Published online by Cambridge University Press

Certainly a value system could be placed on what warrants long-term care, but this also brings forth the question of what data should be created digitally in the first place. Do all field notes and illustrations need to be created digitally? No, they do not, but if these advanced methods are facilitating public access to data, then there is a rationale for doing so. The democratization of archaeological and cultural heritage information is paramount to the future relevance of these fields, and we should be finding better ways to provide access to cultural patrimony, especially when descendant groups have granted scholars permission to do so.[7] If the Internet is to be the means of achieving this, then we should be managing data operability so that we do not distance populations from their patrimony in our attempts to connect them to it.

CONCLUSION

I have argued that workflows in archaeology and cultural heritage management are moving in a direction that is almost exclusively digital, and, just as we are ethically and legally bound to plan for the preservation of our collected and created data, so too are we bound to preserve our digital data. This is a relatively new undertaking and it falls, therefore, to larger archaeological associations to inform members of new ethical and legal stances regarding changes in popular methods, such as the statement recently published by the SAA and tDAR (2013). Independent researchers cannot achieve good practices in data preservation without the aid of preservation services. Thus, preservation services—not archaeologists—are the future stewards of cultural heritage in digital form. If the concept of long-term care is changing, then archaeological associations need to devise a means of evaluation and accreditation for the existing options.

It is remarkable that digital data preservation is underdeveloped as a topic, considering the gravity of the situation and what we stand to lose. Archaeologists and cultural heritage management practitioners work to capture the history of people. Our data are records of lives lived, and the many still-living descendant populations of the cultures we study are an active and receptive audience. To woefully neglect or limit access to these data would be disrespectful and fall short of our obligations to cultural patrimony more broadly. It is thus a cultural imperative that we, as stewards of cultural heritage, focus our attentions on preserving the digital heritage we have already documented prior to capturing more in pursuit of our own academic interests.

Acknowledgments. The research published in this article did not require permits and is the result of work conducted in the Department of Archaeology at Boston University. I am grateful to a number of colleagues and mentors for their input during the various stages of this article, but first and foremost I would like to acknowledge the anonymous reviewers and the editors at *Advances in Archaeological Practice* for their insightful comments. I would especially like to thank Ricardo Elia, Chris Roosevelt, David Carballo, William Saturno, Andrea Berlin, and Sarah Newman for their helpful comments, suggestions, and guidance. Additionally, I am grateful to Kathleen Forste, Guzin Eren, Jared Koller, Paulo Medina, and the rest of the Archaeology Department at Boston University for creating an atmosphere where new ideas can thrive and deep discussions about their solutions are welcome. Finally, I would like to thank ADS, Digital Antiquity, OpenContext, and the other repositories reviewed in this article for continuing to move faster than the rest of us, and for being prepared to be the solutions for problems we have yet to realize.

Data Availability Statement. The data integrated into this article are made accessible through the links provided in the citations. Any comments regarding the presented content may be directed to the author.

REFERENCES CITED

Advisory Council on Historic Preservation
2014 National Historic Preservation Act. Electronic document, http://www.achp.gov/docs/NHPA%20in%20Title%2054%20and%20Conversion%20Table.pdf, accessed July 9, 2015.

Alexander, Patrick H.
2013 Thoughts about Open Access Publishing in a Humanities Context. *Journal of Eastern Mediterranean Archaeology and Heritage Studies* 1(1):97–98.

Altman, Micah, J. Crabtree, D. Donakowski, M. Maynard, A. Pienta, and C. Young
2009 Digital Preservation through Archival Collaboration: The Data Preservation Alliance for the Social Sciences. *The American Archivist* 72(1):170–184.

American Anthropological Association
2012 Resource Development Committee Launches the Anthropology Document Portal. Electronic document, http://www.aaanet.org/membership/Donation/upload/RDC-Anthropology-Document-Portal.pdf, accessed July 10, 2015.

American Cultural Resources Association
1996 American Cultural Resources Association Code of Ethics and Professional Conduct. Electronic document, http://www.acra-crm.org/?CodeofEthics, accessed July 10, 2015.

Archaeological Institute of America (AIA)
1997 AIA Code of Ethics. Electronic document, http://www.archaeological.org/pdfs/AIA_Code_of_EthicsA5S.pdf, accessed April 22, 2015.

Archaeology Data Service (ADS)
2014a ADS Terms and Conditions. Electronic document, http://archaeologydataservice.ac.uk, accessed April 23, 2015.
2014b ADSeasy: About. Electronic document, http://archaeologydataservice.ac.uk/easy/home, http://archaeologydataservice.ac.uk/easy/costing, accessed April 23, 2015.

Archaeological Data Services and Digital Antiquity
2011 Guides to Good Practice. Electronic document, http://guides.archaeologydataservice.ac.uk/, accessed July 9, 2015.

Armstrong, Ann R.E.
2006 Architectural Archives/Archiving Architecture: The Digital ERA. *Art Documentation: Journal of the Art Libraries Society of North America* 25(2):12–17.

Baker, Karen S., and Lynn Yarmey
2009 Data Stewardship: Environmental Data Curation and a Web-of-Repositories. *The International Journal of Digital Curation* 2(4):12–27.

Barton, Justin
2009 3D Laser Scanning and the Conservation of Earthen Architecture: A Case Study at the UNESCO World Heritage Site Merv, Turkmenistan. *World Archaeology* 41:489–504.

Bawaya, Michael
2010 Virtual Archaeologists Recreate Parts of Ancient Worlds. *Science Magazine* 327:140–141.

https://doi.org/10.7183/2326-3768.3.4.313 Published online by Cambridge University Press

Beale, Nicole
2012 How Community Archaeology Can Make Use of Open Data to Achieve Further Its Objectives. *World Archaeology* 44:612–633.

Beaubien, Harriet, B. V. Karas, and William Fitzhugh
2007 Documenting Mongolia's Deer Stones: Application of Three-Dimensional Digital Imaging Technology to Preservation. In *Scientific Research on the Sculptural Arts of Asia: Proceedings of the Third Forbes Symposium at the Freer Gallery of Art*, edited by Janet G. Douglas, Paul Jett, and John Winter, pp. 133–142. Archetype Publications, London.

Berman, Francine
2008 Got Data? A Guide to Data Preservation in the Information Age. *Communications of the ACM* 51(12):50–56.

Bonde, Sheila, and Stephen Houston
2013 Re-Presenting Archaeology. In *Re-presenting the Past: Archaeology through Text and Image*, edited by Sheila Bonde and Stephen Houston, pp. 1–8. Oxbow Books, Oxford.

Borgman, Christine L.
2008 Data, Disciplines, and Scholarly Publishing. *Learned Publishing* 21(1):29–38.

Bruno, Fabio, G. De Sensi, M. L. Luchi, S. Mancuso, and M. Muzzupappa
2010 From 3D Reconstruction to Virtual Reality: A Complete Methodology for Digital Archaeological Exhibition. *Journal of Cultural Heritage* 11:42–49.

Çatalhöyük Research Project
2015 Archive Reports. Electronic document, http://www.catalhoyuk.com/archive_reports/, accessed April 23, 2015.

CLOCKSS Archive
2015 *Controlled* LOCKSS. Electronic document, https://www.clockss.org/clockss/Home, accessed July 12, 2015.

Conway, Esther, David Giaretta, Simon Lambert, and Brian Matthews
2011 Curating Scientific Research Data for the Long Term: A Preservation Analysis Method in Context. *The International Journal of Digital Curation* 2(6):38–52.

Conway, Paul
2010 Preservation in the Age of Google: Digitization, Digital Preservation, and Dilemmas. *The Library Quarterly* 80(1):61–79.

Consultative Committee for Space Data Systems
2002 *Recommendation for Space Data System Practice: Reference Model for and Open Archival Information System (OAIS), Blue Book.* CCSDS Secretariat, Washington, D.C.
2012 *Recommendation for Space Data System Practice: Reference Model for and Open Archival Information System (OAIS), Magenta Book.* CCSDS Secretariat, Washington, D.C.

Creative Commons
2015 About the Licenses. Electronic document, http://creativecommons.org/licenses/, accessed July 10, 2015.

Cultural Heritage Partners
2012 Legal Analysis of Federal Requirements for Curation of Digital Archaeological Documents and Data. Report to the Office of General Counsel, Arizona State University, Tempe, AZ. Electronic document, http://www.digitalantiquity.org/CHPFinal.pdf, accessed July 9, 2015.

CyArk
2003 CyArk: About. Electronic document, http://www.cyark.org/about/, accessed April 23, 2015.
2015 CyArk: Projects. Electronic documents, http://cyark.org/projects/, accessed April 23, 2015.

Dassault Systems
2013 Giza 3D. Electronic document, www.3DS.com/GIZA3D, accessed April 23, 2015.
2015 About 3DS – Dassault Systems. Electronic document, http://www.3ds.com/about-3ds/, accessed April 23, 2015.

Data Archiving and Networking Services
2005 The Data Seal of Approval. Electronic document, http://www.datasealofapproval.org/en/information/about/, accessed April 23, 2015.
2015 License Agreement. Electronic document, www.dans.knaw.nl/en/about/organisation-and-policy/legal-information/DANSlicenceagreementUK.pdf, accessed July 10, 2015.

Dataverse Project
2015 The Dataverse Project. Electronic document, http://dataverse.org, accessed July 10, 2015.

De Reu, Jeroen, P. De Smedt, D. Herremans, M. Van Meirvenne, P. Laloo, and W. De Clercq
2014 On Introducing and Image Based 3D Reconstruction Method in Archaeological Practice. *Journal of Archaeological Science* 41:251–262.

Digital Repository of Ireland
2015 The Digital Repository of Ireland. Electronic document, http://www.dri.ie/, accessed July 12, 2015.

Dublin Core Metadata Initiative
2015 Metadata Basics. Electronic document, http://dublincore.org/metadata-basics/, accessed July 10, 2015.

Dürr, Eugène, Kees van der Meer, Wim Luxemburg, and Ronald Dekker
2008 Dataset Preservation for the Long Term: Results of the DareLux Project. *The International Journal of Digital Curation* 1(3):29–43.

Eve, Stuart
2012 Augmenting Phenomenology: Using Augmented Reality to Aid Archaeological Phenomenology in the Landscape. *Journal of Archaeological Method and Theory* 19:582–600.

Faniel, Ixchel, Eric Kansa, Sarah Whitcher Kansa, Julianna Barrera-Gomez, and Elizabeth Yakel
2013 The Challenges of Digging Data: A Study of Context in Archaeological Data Reuse. Electronic document, http://www.oclc.org/content/dam/research/publications/library/2013/faniel-archae-data.pdf, accessed August 3, 2015.

Favro, Diane
2013 To Be or Not to Be in Past Spaces: Thoughts on Roman Immersive Reconstructions. In *Re-presenting the Past: Archaeology Through Text and Image*, edited by Sheila Bonde and Stephen Houston, pp. 151–168. Oxbow Books, Oxford.

FigShare
2015 FigShare. Electronic document, http://figshare.com, accessed July 12, 2015.

Forte, Maurizio
2014 Archaeology: New Perspectives and Challenges—The Example of Çatalhöyük. *Journal of Eastern Mediterranean Archaeology & Heritage Studies* 2(1):1–29.

Gillings, Mark
2005 The Real, the Virtually Real and the Hyperreal: The Role of VR in Archaeology. In *Envisioning the Past: Archaeology and the Image*, edited by Sam Smiles and Stephanie Moser, pp. 223–239. Blackwell Press, Oxford.

Green, Susie, Andrew Bevan, and Michael Shapland.
2014 A Comparative Assessment of Structure From Motion Methods for Archaeological Research. *Journal of Archaeological Science* 46:173–181.

Guidi, Gabriele, Michele Russo, and David Angheleddu
2014 3D Applications in Archaeology and Cultural Heritage. *Digital Applications in Archaeology and Cultural Heritage* 1: 55–69.

Gunia, Betsy, and Robert J. Sandusky
2010 Designing Metadata for Long-Term Data Preservation: DataONE Case Study. *Proceedings of the American Society for Information Science and Technology* 47(1):1–2.

Hall, Martin
2013 The Emerging Open World. *Journal of Eastern Mediterranean Archaeology & Heritage Studies* 1(1):98–100.

Hanson, Brooks, Andrew Sugden, and Bruce Alberts
2011 Making Data Maximally Available. *Science* 331:649.

Hockx-Yu, Helen, and Gareth Knight
2008 What to Preserve? Significant Properties of Digital Objects. *The International Journal of Digital Curation* 1(3):141–153.

https://doi.org/10.7183/2326-3768.3.4.313 Published online by Cambridge University Press

International Council of Museums
2014 CIDOC-Conceptual Reference Model. Electronic document, http://www.cidoc-crm.org/, accessed July 12, 2015.

International Organization for Standardization (ISO)
2011 Guidelines for Quality and/or Environmental Management Systems Auditing. Electronic document, http://www.iso.org/iso/iso_catalogue/catalogue_ics/catalogue_detail_ics.htm?csnumber=31169, accessed April 22, 2015.

Jeffrey, Stuart
2012 A New Digital Dark Age? Collaborative Web Tools, Social Media and Long-Term Preservation. *World Archaeology* 44:553–570.

Johnson, Matthew
2012 Phenomenological Approaches in Landscape Archaeology. *Annual Review of Anthropology* 41:269–284.

Kansa, Eric
2012 Openness and Archaeology's Information Ecosystem. *World Archaeology* 44:1–23.

Kansa, Eric, and Sarah Whitcher Kansa
2013 Open Archaeology: We All Know That a 14 Is a Sheep: Data Publication and Professionalism in Archaeological Communication. *Journal of Eastern Mediterranean Archaeology and Heritage Studies* 1(1):88–97.

Kansa, Sarah Whitcher, and Eric Kansa
2014 Data Publishing and Archaeology's Information Ecosystem. *Near Eastern Archaeology: Cyber Archaeology* 77:223–227.

King, Gary
2011 Ensuring the Data-Rich Future of the Social Sciences. *Sciences* 331:719–721.

Kintigh, Keith W., Jeffrey H. Altschul, Ann P. Kinzig, W. Frederick Limp, William K. Michener, Jeremy A. Sabloff, Edward J. Hackett, Timothy A. Kohler, Bertram Ludäscher, and Clifford A. Lynch
2015 Cultural Dynamics, Deep Time, and Data: Planning Cyberinfrastructure Investments for Archaeology. *Advances in Archaeological Practice* 3:1–15.

L–P: Archaeology
2015 The ARK Reference Model. Electronic document, http://ark.lparchaeology.com/, accessed July 12, 2015.

Leon Levy Expedition to Ashkelon
2015 Online Database. Electronic document, http://digashkelon.com/current-projects/, accessed April 23, 2015.

Library of Congress
2015 Metadata Encoding and Transmission Standards. Electronic document, http://www.loc.gov/standards/mets/, accessed July 10, 2015.

Mesick, Cassandra
2013 Of Imaging and Imagining: Landscape Reconstruction at Peidras Negras. In *Re-presenting the Past: Archaeology through Text and Image*, edited by Sheila Bonde and Stephen Houston, pp. 65–84. Oxbow Books, Oxford.

Meyer, Élise, Pierre Grussenmeyer, Jean-Pierre Perrin, Anne Durand, and Pierre Drap
2007 A Web Information System for the Management and Dissemination of Cultural Heritage Data. *Journal of Cultural Heritage* 8:396–411.

Molloy, Jennifer C.
2011 The Open Knowledge Foundation: Open Data Means Better Science. *PLOS Biology* 9:1–4.

Museum of Fine Arts Boston
2007 Giza Archives. Electronic document, http://www.gizapyramids.org/static/html/library.jsp, accessed April 23, 2015.

National Archives Records Administration
1998 Trustworthy Repositories Audit and Certification: Criteria and Checklist. Electronic document, http://www.crl.edu/sites/default/files/d6/attachments/pages/trac_0.pdf, accessed April 22, 2015.

National Park Service
2006 [1979] Archaeological Resources Protection Act. Electronic document, http://www.nps.gov/history/local-law/fhpl_archrsrcsprot.pdf, accessed July 9, 2015.

National Science Foundation
2013 Application Policies. Electronic document, http://www.nsf.gov/pubs/policydocs/pappguide/nsf13001/gpg_index.jsp, accessed July 12, 2014.
2015 Public Access Plan. Electronic document, http://www.nsf.gov/pubs/2015/nsf15052/nsf15052.pdf, accessed July 10, 2015.

OpenContext
2015 Data Publishing Guidelines for Contributors. Electronic document, http://opencontext.org/about/publishing, accessed July 10, 2015.

Petrovic, Vid, Aaron Gidding, Tom Wypych, Falko Kuester, Thomas A. DeFanti, and Thomas E. Levy
2011 Dealing with Archaeology's Data Avalanche. *Computer* 44(7):56–60.

Pleiades Project
2014 The Pleiades Project, About. Electronic document, http://pleiades.stoa.org/, accessed July 12, 2015.

Porter, Benjamin W.
2013 Sharing Data Is Hard!—But Worth It. *Journal of Eastern Archaeology and Heritage Studies* 1(1):100–101.

Pratt, Darrin
2013 Not an Either/Or Proposition: Combining Interpretive and Data Publication. *Journal of Eastern Mediterranean Archaeology and Heritage Studies* 1(1):101–102.

Register of Professional Archaeologists (RPA)
1998 Codes and Standards. Electronic document, http://rpanet.org/?page=CodesandStandards, accessed April 22, 2015.

Richards, Julian D., and Judith Winters
2015 Digging into Data: Open Access and Open Data. *Post-Classic Archaeologies* 5:285–298.

Rieger, Oya Y.
2008 *Preservation in the Age of Large-Scale Digitization*. Council on Library and Information Resources, Washington, D.C.

Remondino, F., A. Gruen, J. Von Schwerin, H. Eisenbeiss, A. Rizzi, S. Girardi, M. Sauerbier, and H. Richards-Rissetto
2009 Multi-Sensor 3D Documentation of the Maya Site of Copan. Electronic document http://cipa.icomos.org/fileadmin/template/doc/KYOTO/131-1.pdf, accessed August 3, 2015.

Roosevelt, Christopher H., Peter Cobb, Emanuel Moss, Brandon R. Olson, and Sinan Ünlüsoy
2015 Excavation is Destruction Digitization: Advances in Archaeological Practices. *Journal of Field Archaeology* 40:325–346.

Sanders, Donald H.
2014 Virtual Heritage: Researching and Visualizing the Past in 3D. *Journal of Eastern Mediterranean Archaeology & Heritage Studies* 2(1):30–47.

Shaw, Jonathan
2012 The Humanities, Digitized. *Harvard Magazine* May–June:40–44.

SketchFab
2015 SketchFab. Electronic Document, https://sketchfab.com/, accessed July 12, 2015.

Society for American Archaeology (SAA)
1996 Principles of Archaeological Ethics. Electronic document, http://www.saa.org/AbouttheSociety/PrinciplesofArchaeologicalEthics/tabid/203/Default.aspx, accessed April 22, 2015.
2009 Archaeological Ethics Codes, Charters, and Principles, Electronic document, http://www.saa.org/AbouttheSociety/AnnualMeeting/EthicsBowl/EthicsResources/CodesChartersPrinciples/tabid/199/Default.aspx, accessed July 10, 2015.

https://doi.org/10.7183/2326-3768.3.4.313 Published online by Cambridge University Press

Society for American Archaeology and the Digital Archaeological Record (SAA and tDAR)
2013 Joint Statement: Announcement for Collaborative Program. Electronic Document, http://www.tdar.org/wp-uploads/www.tdar.org//2013/06/jointstatementtdar.pdf, accessed July 9, 2015.

Society for Historical Archaeology (SHA)
2003 Ethics Statement. Electronic Document, http://www.sha.org/index.php/view/page/ethics, accessed April 22, 2015.

The Digital Archaeological Record (tDAR)
2015 The Digital Archaeological Record. Electronic Document, https://www.tdar.org, accessed July 12, 2015.

Visual Past
2015 The Visual Past. Electronic document, http://visiblepast.net/see/, accessed July 12, 2015.

Zenodo
2015 Zenodo. Electronic document, http://zenodo.org/, accessed July 12, 2015.

Zuiderwijk, Anneke, Marijn Janssen, Sunil Choenni, Ronald Meijer, and Roexsana Sheikh Alibaks
2012 Socio-Technical Impediments of Open Data. *Electronic Journal of e-Government* 10(2):156–172.

NOTES

1. "Reuse" in this sense can be both literal, as future scholars might want to use earlier data to ask new questions, and indirect, as future scholars may require a review of the collection strategies and potential biases of the raw data used to form an interpretation. This type of transparency ensures that future scholars can faithfully cite published interpretations.
2. The 3DGiza project was a product of the Museum of Fine Arts, Boston's Giza Archive funded by the Andrew W. Mellon Foundation in 2000 (www.gizapyramids.org). Initially, the archive was created to house digitized copies of George Reisner's 40 years of excavation notes and illustrations, but almost immediately it grew into a georeferenced archive in which users could view notes and illustrations while hovering a mouse cursor over the location of origin within the user interface.
3. Data Archiving and Networked Services (2014) make the following statement regarding the process of contributing to the Commons, "Certain owners wish to permanently relinquish [their] rights to a Work for the purpose of contributing to a commons of creative, cultural and scientific works ("Commons") that the public can reliably and without fear of later claims of infringement build upon, modify, incorporate in other works, reuse and redistribute as freely as possible in any form whatsoever and for any purposes, including without limitation commercial purposes. These owners may contribute to the Commons to promote the ideal of a free culture and the further production of creative, cultural and scientific works, or to gain reputation or greater distribution for their Work in part through the use and efforts of others."
4. Interoperability is made possible by migratory languages such as XML, HTML, or ArchaeoML, which allow data managers or creators to move or migrate digital objects between multiple repositories and archives.
5. Data publishers such as Open Context (2015), Online Cultural and Historical Research Environment (2015), and Zenodo (2015) also have flexible copyright arrangements.
6. Best practices in data preservation were addressed by the Consultative Committee for Space Data Systems (CCSDS 2002) and the International Organization for Standardization (ISO 2011) when they created an ideal organizational structure, or reference model, for a digital archive (CCSDS 2012), called the Open Archival Information System (OAIS) (CCSDS 2012). Unconcerned with the day-to-day practices of data conservation, OAIS (CCSDS 2012) does not validate new and existing archives nor does it take a position on ethical behavior; it is purely a reference model as are ARK (L–P: Archaeology 2015), CIDOC-CRM (International Council of Museums 2014), CLOCKSS (2015), Dataverse (2015), and INVENIO (Zenodo 2015). An archive may state that it follows a reference model, but there is no internal review aimed at reinforcing such protocols. However, the Data Archiving and Networking Systems (2009), an accreditation body, has created a peer-review system that bestows a Data Seal of Approval (DSA) on archive systems that continue to successfully ascribe to their good practices (DANS 2009).
7. It is important to mention that open access to information may not be possible due to issues concerning looting and possibly the wishes of descendant groups. However, publication and preservation firms that are flexible in their access agreements will be able to work with these legal conditions so that the integrity of the data does not suffer for fear of information exploitation. For more information on community and indigenous archaeology in the digital world, please see the August 2015 issue of *Advances in Archaeological Practice*, where digital co-creation is discussed at length.

AUTHOR INFORMATION

Mary Clarke ■ Department of Archaeology, Boston University, 675 Commonwealth Ave., MA 02215 (meclarke@bu.edu)

https://doi.org/10.7183/2326-3768.3.4.313 Published online by Cambridge University Press

Appendix: A Compilation and Comparison of Data Management, Data Publishing, and Data Preservation Repositories Currently Offering Services to Academics

Website	Foci	Licenses	Accepted file formats	Cost	Access	Future sustainability	Guides to best or good practices	Certifications or reference models
Archaeology Data Services (ADS)	Data Management; Data Preservation; Big and Complex Data sets; Interoperability (ArcheoML)	Common Access Agreement; Flexible	Text, Image, Spreadsheet, CAD, Audio, Video, Geospatial, Virtual Reality, GIS, Scanning, Photogrammetry	Free for doctoral students; data management costs are determined by negotiation as are the storage and refreshment costs	Free Membership after signing Usage Agreement	Funded by the UK Arts and Humanities Research Council and Natural Environment Research Council	Yes	DSA; OAIS
ADSeasy	Data Management; Data Preservation; Self Organized Interfaces; Small Data Sets (Max 300 files); Interoperability (ArchaeoML)	Common Access Agreement; Flexible	Text, Image, Spreadsheet, CAD, Audio, Video, Geospatial, Virtual Reality, GIS	Rates are per file and vary depending on the complexity of file format and size	Free Membership after signing Usage Agreement	Funded by the UK Arts and Humanities Research Council and Natural Environment Research Council	Yes	DSA; OAIS
ARK (L-P Archaeology)	Data Management; Data Conversion; Data Sharing; Customizable Interfaces and Overall Structure; Multi-lingual (foreign collaboration); Fully Integrates Spatial Data	None	Text, Image, Spreadsheet, CAD, Audio, Video, Geospatial, 3D Reconstructions	Free to Download and Install; Users program, manage, and curate their own data.	User access can be granted on multiple levels	Maintained by L - P : Archaeology	No	CIDOC-CRM (soon)
Data Archiving and Networked Services (DANS)	Exclusively Dutch Repository for humanities, archaeology, geospatial sciences, and behavioral and social sciences. Interoperable with Netherlands Coalition for Digital Preservation (NCDD)	Open Access; CC0	Text, Image, Spreadsheet, CAD, Audio, Video, GIS, 3D, RDF	Free for Dutch Projects	Open access to users of National Academic Research and Collaborations Information System (NARCIS)	Funded by the Dutch government, and operates under the Royal Netherlands Academy of Arts and Sciences (KNAW) and the Netherlands Organization for Scientific Research (NWO).	Yes	DSA; OAIS; Dataverse
Dataverse	Repository; Promotes Interoperability or linked data to published papers	CC0	Text, Image, and Statistical Documents. More complex data sets will need special arrangements	Free	Free Membership after signing Usage Agreement	Funded by the President and Fellows of Harvard College	Yes	Dataverse
Digital Repository of Ireland (DRI)	Data Repository; Data Preservation; Data Sharing and Reuse; Data Curation (DOI)	CC-BY; Open Access	Text, Image, Spreadsheet, Audio	Free prior to September 2015; Exclusively for Research Organizations of Ireland	Open Access after Terms of Use Acceptance	European Regional Development Fund	Yes	DSA
Federated Archaeological Information Management Systems Project (FAIMS)	Data Management; Data Preservation; Syncs to Android; Mobile in-field Workflows; Analysis and Visualization Tools; Data Archival; Flexible Structure	All Rights Reserved	Text, Image, Spreadsheets, Audio, Video, Geospatial, 3D	Broken Link	Free membership; Data authors determine access to all archived data	Funded by Australian Research Council Linkage Infrastructure Equipment and Facilities Grant (2014–2016)	Yes (ADS and Digital Antiquity)	tDAR
Figshare	Data Repository; Data Preservation (in the future via Portico); Data Publishing; Customizable Structure;	CC0; CC-BY; Flexible for Institutions	Text, Image, Spreadsheet, Geospatial, Audio, Video	Free to $15/ month based on fixed data size packages (max file size is 1GB and max storage size is 20GB); contact for additional needs	Free membership	FigShare LLP	Yes	CLOCKSS

https://doi.org/10.7183/2326-3768.3.4.313 Published online by Cambridge University Press

Website	Foci	Licenses	Accepted file formats	Cost	Access	Future sustainability	Guides to best or good practices	Certifications or reference models
Online Cultural and Historical Research Environment (OCHRE)	Data Management; Data Conversion; Data Sharing; Customizable Interfaces and Overall Structure; Multi-lingual (foreign collaboration); Fully Integrates Spatial Data; Interoperability (wrote ArchaeoML)	Flexible	Text, Image, Spreadsheet, CAD, Audio, Video, Geospatial, 3D, GIS	Scaled Activational and Annual Fees based on project size; Student Discount	User access can be granted on multiple levels	Digital Library Development Center at The University of Chicago	No	OAIS
Open Context	Data Management; Data Publication after Peer Review; Interoperable Data Repository; Data Sharing; Data Citations (DOI); Interoperability	CC0; CC-BY; Flexible	Text, Image, Spreadsheet, Geospatial, Audio, 3D, and GIS. Contact the Editors for specialized datasets (publish@opencontext.org)	Between $500–$6000, depending on the complexity and size.	Open Access	Funded by National Endowment for the Humanities (NEH), The Encyclopedia of Life, the National Science Foundation (NSF), the American Council of Learned Societies, the Alfred P. Sloan Foundation, the Institute of Museum and Library Services, and the William and Flora Hewlett Foundation; Maintained by the Alexandria Archive; and Digital Preservation is maintained by the California Digital Library (CDL)	Yes (ADS and Digital Antiquity)	CIDOC-CRM
The Paleoindian Database of the Americas (PIDBA)	Data Publication; Data Sharing; Exclusively for Paleoindian materials from across the Americas	Open Access	Text, Image, Spreadsheet	Not Specified	Open Access	The University of Tennessee Department of Anthropology	No	Not Specified
Pleiades	Ancient World Geospatial Data Repository; Data Publishing; Data Sharing	CC-BY	All data that provides information regarding historical geographical information about the ancient world	Not Specified	Free Membership after signing Usage Agreement	Funded by the National Endowment for the Humanities (NEH) from 2006 to 2014, the Institute for the Study of the Ancient World (ISAW) at New York University, and Ancient World Mapping Center at UNC Chapel Hill.	No	STOA Consortium
SketchFab	3D Digital Data Publishing; 3D Digital Data Sharing	DMCA; No Rights Reserved	3D File Formats (roughly 30 formats)	Free Basic (50MB/model), PRO at $10/month (200MB/model), BUSINESS at $29/month (500MB/model). Metadata or annotations increase proportionately	Free Membership after signing Usage Agreement	Private Investors	No	No Mention
Swedish National Data Service (SND)	Data Management; Data Preservation; Data Sharing and Reuse	Flexible Copyrights	Text, Image, Spreadsheet, Audio, Video	Made mandatory by EU and is therefore free	Open Access	The University of Gothenburg and the Swedish Research Council	Yes (ADS and Digital Antiquity)	No Mention

https://doi.org/10.7183/2326-3768.3.4.313 Published online by Cambridge University Press

Website	Foci	Licenses	Accepted file formats	Cost	Access	Future sustainability	Guides to best or good practices	Certifications or reference models
the Digital Archaeological Record (tDAR)	Data Management; Data Preservation; Data Sharing and Reuse; Interoperable	CC-BY; Flexible	Text, Image, Spreadsheet, Geospatial, 3D, Virtual	Bulk and per file rates; SAA members get 3 free files per year (30MB max) until 2016	Free Membership after signing Usage Agreement	Digital Antiquity is hosted by Arizona State University and funded by the Andrew W. Mellon Foundation, National Science Foundation, National Endowment for the Humanities, and the Higher Education Funding Council for England of the United Kingdom through the Joint Information System Committee (JISC)	Yes (ADS and Digital Antiquity)	OAIS
Visual Past	Data Collection; Data Management; Data Publication; Integrated Analogue and Virtual Content for Public Interaction and Tourism	Open Access	Text, Image, Spreadsheet, Geospatial, 3D, Virtual	No Mention	Open Access	No Mention	No	Open Context Platform
Zenodo (formerly OpenAIRE Orphan Record Repository)	Data Management; Data Preservation without the guarantee of long term viability (see policies); Data Sharing and Reuse; Data Curation; Data Citation (DOI); Interoperable with Dropbox, Github, among others	CC0; CC-BY; CC-NC; CC-ND; Flexible Copyrights	All Formats	Free for modest research (max 2GB/file), but contact them directly for the costs of larger datasets (info@zenodo.org)	Open Access	European Commission; OpenAIRE; CERN	No	INVENIO (CERN)

https://doi.org/10.7183/2326-3768.3.4.313 Published online by Cambridge University Press