

Who Cares About 3D Preservation?

Jennifer Moore¹, Hannah Scates Kettler²

Abstract

The preservation of 3D research data is a present and emerging need. An increasing number of researchers are generating, capturing, and/or analyzing 3D data, but they rarely focus on preservation or reuse of that data. This paper and presentation describe models of 3D data creation and use, outline the specific concerns for this data type, unpack the complexities and challenges of preserving it, and examine existing 3D data preservation resources while working through local case studies from the field of anthropology. Directions on how to move digital 3D data preservation forward will be discussed.

Keywords

3D digital data preservation; anthropology; archaeology, digital humanities

Introduction

Anthropologists are concerned with both physical and digital 3D data. In some cases, they are integrating digital 3D capture and/or creation into workflows as a preservation mechanism for artifacts, faunal remains, casts, sites, etc. However, as with other digital data types, creating a digital record of an object or place does not provide stable preservation for said artifact unless the digital data are also treated for preservation.

Digital preservation of data is widely accepted as necessary to protect data against loss and obsolescence, particularly 'where the data are non-reproducible or extremely valuable'. (DCC, n.d.) Digital preservation ensures that data are in an appropriate state for long-term access and reuse. Inserting preservation actions and behaviors into any type of research data workflow can be tedious, as the process requires forethought and is best accomplished when built into the methodology of a project. In actuality, preservation often occurs as an afterthought, at the end of a project, at which point the preservation of the data becomes more difficult if the data have not been appropriately prepared and administered throughout the project lifecycle. Digital preservation principles can be applied more easily to some types of data than others. 3D data fall within the latter category and require specific actions that may not fit typical curation treatments.

Digital 3D data preservation is a burgeoning topic among practitioners and data curators alike. As with all data, digital 3D data preservation actions would be most successful if workflows were established using best practices and standards at the outset, but this has proven to be much easier said than done. Creating and capturing digital 3D data requires intense research and skill development in order to accomplish immediate goals, so the added complication of long-term data stability is typically not at the forefront of methodology considerations. While there is no shortage of literature describing capture and creation methods, the lack of consensus on how to *do* digital 3D preservation for the long-term is a major barrier. The best practices that exist are not yet expansive enough for adoption. Research shows that institutions that are implementing any kind of preservation actions are often doing so *ad hoc*. The authors of this paper recently conducted a survey³ of practitioners and curators, which indicated there is a need and a great deal of support for collaboratively developed best practices and standards.

Literature Review

As early as 2009, Julie Doyle wrote about the need for digital 3D preservation standards, asserting that establishing principles of authenticity and reuse is of the utmost importance. She estimated that the most important step to be taken to ensure long-term preservation is setting up workflows for emulation and metadata creation (Doyle, 2009). More recently, the Archaeological Data Service (ADS)⁴ produced a case study documented in *Curating Research Data, A Handbook of Current Practice* (Johnston, 2016), which briefly outlines some 3D data curation concerns. The case study suggests that preservation of reusable 3D objects is challenging because projects are often focused on an end product. The ADS & Digital Antiquity, the London Charter, and a group called 3D-Icons⁵ have put forth some limited recommendations that offer a foundation but are also purposefully vague or incomplete, making the preservation landscape difficult to navigate.

The London Charter⁶ was developed in 2006 ‘as a means of ensuring the methodological rigor of computer-based visualization as a means of researching and communicating cultural heritage. Also sought was a means of achieving widespread recognition for this method’ (London Charter, n.d.). It outlines six basic objectives, which are to create standards that ensure methods are rigorous, intelligible, useful, sustainable, and extensible by the community of practice. The objectives are manifested in the practice of six basic principles. Of the principles, 1) implementation 2) aims and methods 3) research sources 4) documentation 5) sustainability and 6) access, aims and methods as well as documentation and sustainability are particularly relevant to data preservation. While these principles provide a valuable framework for considering necessary workflows at a high level, no suggestion of specific recommendations are made.

The ADS and Digital Antiquity⁷ collaborated to create the *Guides to Good Practice* to ‘ensure digital data access and long-term preservation (*Guides To Good Practice*, n.d.). The guides cover a multitude of methods and data types in use by archaeologists, including close-range photogrammetry and 3D laser scanning. The guides also offer information related to formats, approaches to documentation, and suggest minimum, file-level metadata elements and workflows, but they remain short of standards and are far from exhaustive.

The mission of 3D Icons is to create ‘highly accurate’ 3D models of important heritage sites in Europe. 3D-Icons produced a report that focused on adapting the CARARE schema⁸ for 3D models. The aim is to assure quality through providing provenance, transformation, and paradata description. The CARARE schema was developed for describing digital items for cultural heritage organizations in Europe (D’Andrea & Fernie, 2013). The work done by 3D-Icons, though also narrow in its focus on provenance and paradata, could be useful in building more widely accepted digital 3D data metadata standards.

Methods of 3D Data Creation and Use

There are four ways in which digital 3D data are created: through free-form, physical, real-world measurements (i.e. with a tape measure); algorithmically, as with the 3D scanning methods discussed later; observationally, and by using treatises or historical research. Many methodologies use a mixture of these creation methods.

Data created through the free-form method can be an expression or visualization aid for concepts, or a representation of real-world space. Free-form 3D modeling refers to the lack of any material that

represents the object or space being modeled. Much of this type of work is being done in virtual reality spheres. This process allows for some artistic license in 3D representation and some fluidity in regard to the requirements for 3D preservation. Augmented Asbury Park⁹ is a representation of lost historical space. The project represents the heyday of the Asbury Park boardwalk and places the missing buildings and attractions in their physical space using augmented reality. A guided tour is available in the form of a physical booklet one can carry around the park, using coded images to bring in 3D models of some of the boardwalk attractions by means of an augmented reality application. 3D models were created to represent the buildings, which in many instances were created free form and without digital guides. The models are not 3D scans but are representations and interpretations of space based on photography of various similar objects and the allowances of the physical space at Asbury Park.

In this instance, there is much more flexibility in terms of what is necessary to preserve. It may be more of a prerogative for the data preservation specialist and researcher to preserve the raw data, the experience of the data, or the presentation of the data. Continuing with the example of Augmented Asbury Park, it may make more sense to preserve the physical booklet and virtual 3D object relationship which would require that the 3D data be preserved in a way that can allow for display on a new platform or emulation in conjunction with the physical booklet artifact. As with many other projects, and especially digital humanities projects, the physical and the digital artifacts are inseparable and should maintain their relationship to support the research being conveyed.

3D models based on the physical process of measurement are, for the purposes of this article, distinct from algorithmically generated 3D models because the *act* of measurement is not a proprietary process and can easily be reproduced by anyone with a ruler. 3D data creation based on physical measurements requires one to go out into the field or handle the specific object and take detailed measurements that can be visualized using 3D modeling software. The technique requires extreme attention to detail as well as the ability to maintain accurate notes that would represent the analog of the 3D model. In a lot of these cases, the 3D model is the scientific redundancy (which has heretofore been missing) in fields like archaeology. A 3D model based on physical measurements is a much more transparent form of 3D data creation. The use of 3D scanning and Geographical Information Systems (GIS) in archaeology has expanded the possibilities of 3D modeling over much larger areas but is much more dependent on computationally generated measurement. The automation of 3D data creation places the processes of measurement in a 'black box' and makes the processes proprietary and obscure.

Algorithmically generated 3D data, like the 3D scanning and GIS methods mentioned above, are much harder to evaluate, reproduce, and decouple from the technology used to create them. 3D scans measure distance, angle, reflectance, and color, depending on the type of scanning method one uses, such as computerized tomography (CT) scanning, laser, structured light scanning, and photogrammetry, to name a few. The processes used to evaluate said measurements are not open sourced, so that a researcher hoping to reproduce the data creation process can evaluate the integrity of the data. At first, this may not seem to be an issue for preservation, but upon further reflection it is integral to any 3D preservation process. When data need to be reproduced or migrated for preservation purposes, there is no way to evaluate whether the reproduction is faithful or the migration successful without a clear understanding of how the data were created in the first place. Since much of today's 3D data streams from scanners is entirely proprietary, the likelihood of successful reproducibility of 3D scanning data is low to non-existent.

Take, for example, the laser-scanned 3D Collection of Artifacts¹⁰. These data represent many hours of labor and are specimens used for anthropological research. These scans are, in many ways, digital surrogates for the physical specimens. Lasers accurately measured and reproduced the objects at a level of faithfulness researchers were satisfied with, but the process that occurs during scanning and computational interpretation is entirely proprietary. This process has its own uniquely coded

methodology and software, which generates the end product that is available on the Web. In several years, these data will need migration to new servers and a new platform and will need new standards of preservation. Although 3D scanning methods are popular, they produce a singular representation of raw 3D data that makes articulating the importance of data preservation much more difficult because of its opaqueness and unverified processes and methodologies. To mediate this, one could argue for the usage of open-sourced 3D scanning technologies to facilitate the preservation of their outputs and bolster the integrity of the generated 3D models as reproducible, or at the very least, transparent research objects.

Not all 3D data are based entirely on the presence of physical artifacts. A 3D data creator could rely on traditional research methods to determine how to represent artifacts in virtual space. This method of 3D data creation has much in common with free-form modeling. 3D data may be based on historical research and interpretation like any other work one would publish. The 3D data are imbued with thoughts, decisions, assumptions, and technological or expertise limitations, and may represent a range of artifact types, including imaginary, lost, or degraded objects. The 3D model represents a scholarly argument in much the same way as a scholarly monograph. The resulting 3D data therefore requires the same level of preservation as an article or book, with the same attention paid to reusability and shelf life.

The long-running Digital Hardian's Villa Project¹¹ created 3D data that represents not only the empirical archaeological data, but also contextual research about the villa. The researchers have used a combination of free-form, physical measuring and scanning to extrapolate Hardian's villa footprint into a fully three-dimensional space. A website has been built around the research involved in the creation of the 3D data and houses not only the models, but also the bibliography and citations necessary to its creation, the paradata describing the decision processes, the photos on which the 3D model is based, the videos of the space, and interviews with specialists discussing what they know about the individual spaces. All of this contextualized research is underpinned by the process of 3D data creation as much as the contextualizing research bolsters the resulting 3D model. The 3D data are a part of a feedback loop with all other data types in this digital monograph and should be rendered as stable as other scholarly outputs.

To complicate 3D data creation, much 3D work is done using a combination of the methods mentioned thus far. 3D data represents the gambit of individualized, single processing work (as with 3D scanning) but may be as complex as a 3D thesis or publication that combines 3D scanning, free-form modeling, data visualization, and traditional research methodologies.

Unpacking Digital 3D Data Capture

Data can be captured to create 3D models in various ways, including close-range photogrammetry, computed tomography (CT) scanning, structured light scanning, laser (time of flight, phase shift, triangulation) scanning, etc. To describe all of the methods of collection is beyond the scope of this paper, which focuses on some of the methods commonly employed by anthropologists. The methods include close-range photogrammetry, structured light scanning, and triangulated laser scanning.

Close-range photogrammetry is a process by which a camera is calibrated to capture multiple images of an object. Many types of cameras can be used to accomplish this task, but higher-quality images enable more reliable 3D reproductions. Calibration is specific to the lens and camera used; this calibration accounts for any distortion by the lens, sensor, or processor. An external control may be added to define the data and/or geometric constraints (Guides to Good Practice, n.d.). The images are stitched together and processed by software that will then allow for a number of outputs, including a point cloud, 3D polylines, mesh, or raster graphics.

A point cloud can be converted to RCS/RCP files and attached to a CAD drawing to create a 3D CAD model, or its output(s) can be manipulated in other 3D software such as MeshLab¹² (Martorelli, 2014). An example of this method is illustrated by the work of Martorelli, Pensa, and Speranza in their paper 'Digital Photogrammetry for Documentation of Maritime Heritage'. They presented three case studies that use photogrammetry to analyze historical boats: Tomahawk, Refola, and Nada. The case study of the Tomahawk described the methods, which included calibration, marker placement, image capture, processing, point cloud generation, and 3D model analysis. The authors outlined the process for creating the 3D model as a cycle of 'correction, surface layout, sectioning, further correction...'. (Martorelli, 2014). If information on calibration and processing was recorded, it was not offered in this article.

It should be noted that 3D digital data preservation was not the purpose of this research. By the authors' own estimation, the method of photogrammetry did not result in the greatest accuracy, but they determined it was a good fit because it was a 'quick and inexpensive' way to acquire the data they wanted. This example inspires many questions. Without detailed documentation of the methods, are these data reusable? With iterative processing of the model, at what stage should it be preserved? Are these data suitable for preservation, given that they are probably not accurate? If so, in which format(s) should they be preserved? Would they meet selection criteria for preservation, assuming that such a criteria existed?

Triangular laser scanning is ideal for close-range scanning. As the name suggests, this type of scanning uses triangulation to collect precise point locations by projecting a laser line onto an object, which bounces back to the sensor. Triangular laser scanning is also dependent on accurate calibration; some scanners can collect highly accurate data, and some have color capture capabilities. Molloy et. al (2016) describe methods of using triangular laser scanning to investigate the edge wear on prehistoric tools to understand their function. The authors describe establishing resolution by setting the step rate of the laser, which determines the point-to-point distance. Denser point clouds produce higher resolution. The project is said to have used the highest resolution settings, but specifics are not listed. Eight to 12 static positions were rotated to capture the entirety of each object, and the software stitched the various scans together. It's noted that once the scan finished producing a point cloud, the superfluous data were removed, which the software documented. Molloy et. al were satisfied with the results of the laser scanning and commented on the portability and efficiency of the process.

This research project also employed photogrammetry by capturing raw images at high resolution, processing, and post-processing to remove noise. By the authors' assessment, the laser scanner excelled in capturing small details but was weaker in capturing sharp changes of direction, reflective surfaces, and damage to the edges of objects. In one case, the scanner was unable to capture the edge data, and, while either side of the blade was captured, there was a data gap that could only be completed by large-scale hole filling, which leads to inaccurate data. Each method the authors employed had weaknesses, but it's noted that they intended to combine methods for a more complete model. Questions of what to preserve are again raised by this example, due to the iterative workflow, the problem of processing, corrective measures (like hole filling), and the combining of data from disparate methods to make a complete model. This project seems to have recorded some essential information for the creation of useful metadata, although it was not offered as a part of the paper.

Structured light scanning is accomplished by projecting white light onto a surface to display a series of organized patterns (Wachowiak, 2009). Cameras record the changes to display a surface, and the software algorithmically calculates the distances. Structured light at close range is good for capturing a small point-to-point distance or resolution. Calibration of the light and camera is essential for accurate 3D reproduction. Because the scanner only captures what it can see, a series of scans must

be completed, often using a rotary table (for small objects) and changing the angle of the object to cover all of the area. This means that overlapping data are collected, and that the camera often collects noise from background interference, etc. Laura Niven (2009) used structured light scanning to create a reference collection of faunal skeletons for zoo archaeologists. Researchers in this example spent four years developing a workflow, which is well documented in the article, with details such as specimen size ranges and lens choice. The structured light scanner they used allowed for capture with or without color. Their scanning took place in a photographic light tent on a black background, which reduced shadow and noise and made processing more efficient. The specimen was scanned in 3-4 rotation steps. The software stitched the scans together after two steps, and alignment was automatically added throughout the process after the first two steps.

Following the first set of rotations, the specimen was repositioned, and the new set was scanned and aligned. Once all of the scans were completed, the data were merged into a final scan, with overlap automatically adjusted. The outputs of this project were .ply (polygon file format, stores multiple properties) files for 'archival purposes', as well as .stl (stereolithography, basic surface description), and .obj (simple geometry), which were used to create 3D PDFs as well as some 2D image formats. This example was very detailed in local documentation and took preservation into account, but the problem of processing and the subsequent questions of when and what needs to be preserved still exists (what are the 'raw' data?), even in this careful example.

Local Case Studies

Fluxus Digital Collection

The Fluxus Digital Collection was a Digital Humanities collaboration between a faculty member, technologists in Information Technology, and the Digital Humanities Librarian at the University of Iowa. The intention was to create an online contextualizing exhibit of the Fluxus West art collection housed in the University of Iowa Special Collections. Grappling with the library's charge to preserve, the scholar's prerogative to study, and the artists' intent to make art interactive and accessible, the collaborators decided to reconcile these issues and create 3D models of select objects as a form of preservation and as a mode of interactivity for patrons.

In 2012, the University of Iowa did not have 3D scanning equipment or the expertise to create 3D models from scans. As an example and a consult, the University of Iowa Libraries used the Sousa Archive 3D flute collection at the University of Illinois at Urbana-Champaign, purchasing the same Strata 3D (photogrammetry) Scanning setup and hiring a 3D modeler to scan and post-process the materials.

The Fluxus project was, in part, a multi-departmental library endeavor to test the potential for 3D scanning as a preservation tool for fragile Special Collections objects. As part of this endeavor, the University of Iowa Libraries created a bespoke 3D workflow and digitization effort involving various units: Special Collections organized and pulled materials; Preservation & Conservation assessed the stability of the materials and made recommendations on handling; the Digitization department oversaw the standard scanning digitization of the Fluxus West Collection; and the *né* Digital Research & Publishing department was in charge of figuring out what a 3D scanning workflow would look like and how it would fit into more traditional digitization efforts. This was not an effort to test the feasibility of 3D preservation but to test whether 3D scanning could be used as a form of preservation of Special Collection objects. This project, like many, was focused on the outcome of 3D scans and not proportionally concerned with the preservation of said scans (the departure of the Digital Preservation Librarian during the project added another hurdle to the initiative).

The creation process was documented for those wishing to recreate the scanning workflow, as was information contextualizing the 3D scanning data. The workflow that resulted from this project was one that complemented the more traditional 2D scanning workflow, and the resulting 3D scans were treated similarly. The London Charter was used as a guide for best practices for the visualization, but it was deemed insufficient for a library implementation focused on 3D preservation. All objects were barcoded and identified in the digital catalogue, renamed and described, and the 3D scanning folders reflected the established file naming conventions used for all other catalogued Special Collection objects, folders, and boxes. Once the objects were photographed and named in accordance with the institutional naming conventions, the post-processing of the 3D scans occurred.

Most of the 3D data cleaning occurs in the post-processing phase of a 3D scan. The scan goes through various types of transformations that corrupt the raw data but also support end-user needs and reflect the researcher's intended portrayal. The resulting augmented data represent a new digital object that is in some cases rather loosely derived from the raw data collected during the scanning process. For the Fluxus project, there was an interesting tension between the project as a projection of an intentionally ephemeral art movement, the necessary steps to create digital representations of the artifacts, and the motivation of the library to preserve that which was meant to defy cataloging and preservation. For the researchers, the creation of the 3D scans was an interesting aspect of the collection and was worth documenting and keeping for the long term.

Unlike other projects, where keeping 'everything' is more of a reflexive process, with the Fluxus project, the need to preserve both the final 3D dataset and the raw data -- including the photography images, original 3D datafiles (in the ISO standard of stereo lithography (.stl) file format), and the proprietary Strata/3DSOM software files -- was clearly articulated. Due to a lack of guidance on the subject, the library was unsure what 3D data to keep and what would be useful in the future if the University of Iowa 3D scanning processes needed to be replicated. The preservation process did not include preserving the mode of display of the data (i.e. the now-antiquated Adobe Flash player), nor the interaction of the user and the 3D object.

The space required to house all of these data was significant, especially compared to other digitization efforts. After a reevaluation of the project scope, it was decided to include only a handful of 3D scans rather than a comprehensive 3D library as was originally intended. The space needed to preserve the various products of the 3D scanning was as unanticipated as the lack of guidance regarding the preservation of 3D data. Additionally, the Fluxus project coincided with the decline of Flash as a standard for presenting multimedia objects (including 3D) online (Howard, 2012), and when WebGL (Web Graphics Library) was just emerging as a possible replacement. Packaged with the 3D object is the README file that captures the workflows taken, the contextualizing paradata about the files and objects within the folder, and a list of potential exports and uses of the data.

The outcome of this project was an evaluation and set of recommendations that remain internal to the University of Iowa Libraries regarding the workflow of 3D scanning as part of a digitization effort for Special Collection materials, as well as a set of recommendations moving forward regarding 3D as a potential form of preservation. As with many such projects, the documentation and recommendation remains relatively project-specific and internal to the institution. This lack of transparency is not uncommon, as evidenced by a recent survey of 3D practitioners and information specialists, but has the unintended consequence of perpetuating the *ad hoc* and esoteric nature of workflow and standards development for 3D preservation.

Digital Baboon

Digital Baboon is an overarching title for the ongoing digital representation of analog data collected by anthropology professor Jane Phillips-Conroy of Washington University in St. Louis and anthropologist Cliff Jolly for their Awash National Park Baboon Research Project in Ethiopia. These data, which span a 30-year period, are being digitized, managed, and preserved for the long-term. Analog data types include field sheets, slides, palm prints, and tooth casts.¹ Items are being scanned in 2D apart from the tooth casts. Plaster casts were taken in the field from captured baboons and correspond with their other data, all of which were collected to determine the hybridity of the olive and hamadryas baboons from the study area's hybrid zone. The tooth casts themselves are important for showing tooth wear, which is enhanced by the ability to scale the 3D models.

More than 2000 casts were collected in the field. Due to the fragile nature of the casts, it is imperative that they are scanned and preserved. To that end, the securely wrapped plaster casts are transferred to the 3D scanning lab, and each cast is given a year and ID number that is related to the capture ID embedded in the file naming convention. Each item is scanned with an HDI Advance R3X¹³ structured light scanner, which is calibrated regularly with a calibration board. The calibration attempts to achieve 80% coverage (at minimum); to achieve this calibration, the scanner must often do around 100 captures of the calibration board. Locally developed documentation is recorded as a part of the technical metadata in a spreadsheet along the following additional attributes: date scanned, by whom, cast number, cast filename, resolution, coverage, average reprojection error, disparity error, point distance, smoothing, hole filling, editing other, file types generated.

The scans are completed using a rotary table with 12-18 rotation stops. The scan is repositioned when the rotation set completes. Each set of rotations align automatically, and the technician manually puts the sets together and finalizes the scan, removes the noise and fills only very small holes. The files are exported to ascii, .obj., and .stl. To date, they have not been exported to .ply, but all of the original project data has been kept, and an export to .ply may be added to the workflow. To assure that the scans are accurate, a sampling of them is measured using the software against the physical object using calipers. In addition to machine-generated metadata, descriptive readme files are created for the overall collection and subcollections, which include calibration details. Scans are moved to Washington University Libraries' servers and regularly backed up to the university's servers. Currently, there is not a public access point for the information, and so the data are not accessible to anyone outside the project. While the methodology for this project has been carefully thought through, guidelines for resolution, processing, and formats and standards for metadata were sought to no end. Left without definitive answers, localized practices were implemented.

DCN Review

Funded by the Alfred P. Sloan Foundation, the Data Curation Network (DCN) is an initiative that enhances data curation services for participating university libraries. Part of the mission of DCN is to evaluate and improve skills, workflows, and best practices for data curators. To this end, DCN ran a pilot in the fall of 2016 asking data curation reviewers to evaluate submissions of supplementary data and metadata to the Data Repository for the University of Minnesota. The submission of interest to 3D preservation was *Supplementary Data for Reconstructing Past Craft Networks: A case study using 3D scans of Late Bronze Age swords to reconstruct specialized craft networks* (Golubiewski, 2016). Careful review of the supplementary data and metadata found that the methods by which the submitted 3D data were collected were not fully documented in the metadata. The data consisted of

.bmp files, and the metadata included documentation on necessary software for viewing. Presumably, using methods of photogrammetry, these images could be stitched together. However, because the methods of data collection were not clearly described in the metadata, it was necessary to review the dissertation to understand the provenance of these .bmp files.

The dissertation included clearer documentation; the methodology included scanning the objects with a structured light scanner, creating screen captures from those, and analyzing the shapes. The gaps in the metadata make reproducibility and reusability unlikely, because curators cannot preserve data that lacks the technical metadata needed to inform future users of what they are working with and how it came to be. This case demonstrates that, without standards and best practices, there is no criteria for researchers to report methods or reliable rubric by which data curators can assess whether metadata are complete.

Iowa City Archaeological Data-Dump

In fall 2016, a group of researchers from the University of Iowa's Department of Geographical and Sustainability Sciences approached the library about housing their LiDAR (Light Detection And Ranging) data for public dissemination. This large dataset consists of many million vertices point clouds, each constituting a gigabyte to several gigabytes each. These data are also in a proprietary file format that is consistent across LiDAR data creators but is not a standard that the library world considers sustainable (Digital Preservation Handbook, 2017).

This data, however, represents the *de facto* data of an archaeological site that is, at present, inaccessible to the public. The researchers intend to make the data available online and to preserve it "long-term" for public use. The conundrum, insofar as the library is concerned, is determining what file format the data should be moved into, so that it can be 1) disseminated to the public 2) preserved for posterity as archaeological data, and 3) shared with an appropriate level of degradation, so that the files are not large and unwieldy using current technologies. Deciding which data are the most appropriate for preservation is a key issue.

Using the model set forth by the Fluxus project, the University of Iowa Libraries could commit to preserving everything – original files and all derivatives. At this point, the library does not know the state of the data which, as an institutional repository, they are required to preserve. Likely, equipment metadata are sufficient, as they are typically created via an automated process like the aforementioned scanning project case studies, but the archaeological and anthropological data surrounding the scan data are most likely recorded elsewhere. Additionally, this project will require that the libraries expand their definition of 3D preservation to include that of the user interface to the data. At present, the project will include a web interface for the 3D data using a Javascript library called 3DHOP¹⁴ and build upon this interface to include data points not present in the raw 3D LiDAR data.

Pending any upcoming recommendations and standards set by the library and archives community in coordination and conjunction with subject specialists, this project has the potential to be as *ad hoc* and bespoke as the ones listed thus far. With varying degrees of contextualizing documentation (for example a README.txt file packaged with all the 3D data outlining the purpose, file types, and decision processes), and the potential for recreation of bespoke metadata schemas, the utility of the 3D LiDAR data preservation practices is suspect. The idiosyncratic creation of preservation practices has the potential to compound the technological barriers of sharing proprietary 3D data, so that one cannot find nor interface with other 3D datasets, hampering (or making impossible) new discoveries and 3D research.

Questions for the Community

These case studies demonstrate that coordination is needed between stakeholders, including data collectors (researchers) and data curators (archivists and librarians), to create best practices and standards that reflect the needs of everyone involved and foster digital preservation within and between institutions. Such a community could collectively answer questions such as:

What should be preserved? This is not only a question of selecting valuable, important, rare, or interesting data. Other basic questions include: at what stage is a digital object most stable, and what does raw data mean in the context of a 3D model? Because of the often-iterative nature of the scanning process -- one 3D model is comprised of many scans and/or other 3D models -- levels of processing, correction, and combination should be considered. Further, there is the question of format: what is the most stable, authentic, and reusable format in which data should be saved? The challenge of storage space for relatively large objects is real, especially in a world where curators are cautious of destroying any type of 3D data. For even a small item, a project folder consisting of raw data and exports to .stl, .obj, and .ply formats can be relatively large for each object (e.g., 10gb) compared to other digital objects, like scanned manuscript materials.

What metadata is necessary? Integral to the preservation of the generated 3D data is the description and cataloging of the data to encourage discoverability and provide context to projects like the ones mentioned thus far. Many methodologies, as well as software and hardware, can be used to create 3D models. What level of description is necessary to effectively document the creation, augmentation, and use of 3D data for future research endeavors? How would metadata look different for 'raw data' verses the final product for end use? What metadata is necessary to help the utility of the data? What metadata would facilitate emulation or the reproducibility of the 3D data?

What is the role of emulation in 3D data preservation? Open-source options exist for opening and viewing 3D model exports. However, working with 3D data in raw format is often done with proprietary software. The outputs of 3D work, therefore, are proprietary and do not directly translate to data that have the potential to be reused outside of emulation of the software in which they were created. In addition to questions around what constitutes raw data, concerns exist about the software in which the data are created and the preservation of the workflows used to create the data output. This is especially important when one is attempting to recreate the 3D data or replicate the 3D scanning process. Coupled with the preservation of workflows or output processes is whether the software (again, typically proprietary) can be maintained alongside the raw 3D data. The reason for this is the maintenance of the 3D creation process as well as the output's presentation and dissemination mechanism, potentially reaching into the possibilities of preserving user experience as a component in need of preservation consideration. Should proprietary software be emulated? Is emulation of a proprietary software prohibited? If it is not possible to emulate it, how can the original raw project data be protected from obsolescence? What kinds of commitment and coordination are required between practitioners, curators, and commercial vendors?

The projects mentioned here represent only a few of the possibilities and idiosyncrasies that occur while working with digital 3D data. Whether the data are created by free-form modeling practices or algorithmically generated as part of a 3D scanning project, without best practices and recommendations for 3D data preservation, the 3D data being captured in this era will be lost to time. Such preservation standards are essential for building capacity in libraries, other repositories, and archives to accommodate the data long-term. That includes considerations of what exactly constitutes raw data, the outputs that result from the 3D data use, and the software and equipment used to

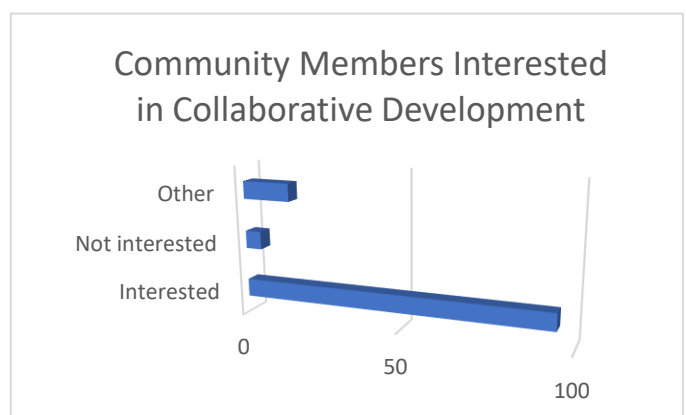
create and disseminate the data. Lack of standards is a detriment to the advancement of 3D scholarship as a whole.

This paper has mentioned several different types of raw 3D data, all distinct and individually decided upon by the project members and not necessarily by some consensus born out of practice. At this point, the aforementioned project members are attempting to predict the possible future use and reuse of their data. As 3D projects extend beyond the lifespan of the technology, concern about preservation and migration increases. Questions are arising about use and reuse of raw 3D data, and about which types of raw data will be useful in future. Does the original, unaltered data represent the raw data that will be most useful in future? Or is the priority the preservation of the 3D data that represents the intended outcome, i.e. the published data?

Moving Forward

An essential element that was present directly or implicitly in every example in this paper is metadata. Developing metadata standards specifically for 3D objects has great potential for growing 3D data preservation. Documentation and metadata are key to creating data that are reusable well into the future. Good documentation and metadata standards, which a set of best practices would address, represent the waterfall impact of detailing data and processes that can inform and bolster the efforts of processes, software emulation, data migration, and dissemination. Creating a standard for metadata and a set of best practice recommendations would have immense impact on the overall preservation and interoperability of 3D research.

Conversations are occurring at conferences¹⁵ and on listservs¹⁶ about the problem of preserving digital 3D data. Efforts are underway to set the foundation for coordinated efforts regarding 3D preservation between stakeholders. The authors of this paper are involved in ongoing conversations regarding 3D support focused on the practices of data creators. These conversations typically focus on outlining the different ways in which 3D data are created by scholars, but little attention is given to preservation, mainly because there are too few preservation specialists in the room.



The authors have since collaborated in administering a survey¹⁷ regarding preservation practices, which revealed that many invested library and archive professionals are thinking about the problem of 3D data but have yet to come together with data creators to figure out what is needed to support their research. The next step is pulling the information from these previous conversations and working toward solutions together.

References

- THE DATA CURATION NETWORK: DISCUSSING AND PROVIDING SOLUTIONS FOR DATA ISSUES IN A COLLABORATIVE WAY | AGRICULTURAL INFORMATION MANAGEMENT STANDARDS (AIMS). (n.d.). Available from: <http://aims.fao.org/activity/blog/data-curation-network-discussing-and-providing-solutions-data-issues-collaborative-way>
- DOYLE, J., VIKTOR, H., & PAQUET, E. (2009). Long-term digital preservation: preserving authenticity and usability of 3-D data. *International Journal on Digital Libraries*, 10(1), 33–47. Available from: <https://doi.org/10.1007/s00799-009-0051-7>
- GOLUBIEWSKI-DAVIS, K. (2016). Reconstructing Past Craft Networks: A Case Study using 3D scans of Late Bronze Age Swords to reconstruct Specialized Craft Networks. Accessed Available from: <http://conservancy.umn.edu/handle/11299/181730>
- GROSMAN, L., SMIKT, O., & SMILANSKY, U. (2008). On the application of 3-D scanning technology for the documentation and typology of lithic artifacts. *Journal of Archaeological Science*, 35(12), 3101–3110. Available from: <https://doi.org/10.1016/j.jas.2008.06.011>
- GUIDES TO GOOD PRACTICE: 3D_TOC. (n.d.). Available from: <http://guides.archaeologydataservice.ac.uk/g2gp/Contents>. [Accessed May 14, 2018]
- GUIDES TO GOOD PRACTICE: PHOTOGRAM_1-1. (n.d.). Available from: http://guides.archaeologydataservice.ac.uk/g2gp/Photogram_1-1. [Accessed March 24, 2017]
- HOWARD, BILL. (2012). *Flash – Chrome for Android Beta*. Retrieved November 20, 2017, from Adobe AIR and Adobe Flash Player Team Blog: blogs.adobe.com/flashplayer/2012/02/flash-chrome-for-android-beta.html
- DIGITAL PRESERVATION HANDBOOK (2017). Available from: <http://www.dpconline.org/handbook/technical-solutions-and-tools/file-formats-and-standards>. [Accessed November 20, 2017]
- DIGITALE 3D REKONSTRUKTIONEN IN VIRTUELLEN FORSCHUNGSUMGEBUNGEN. (n.d.). Available from: <https://www.herder-institut.de/forschung-projekte/laufende-projekte/digitale-3d-rekonstruktionen-in-virtuellen-forschungsumgebungen.html>. [Accessed March 24, 2017]
- LONDON CHARTER. (n.d.). Available from: <http://www.londoncharter.org/>. [Accessed March 24, 2017]
- CIDOC CRM. (n.d.). Available from: <http://www.cidoc-crm.org/>. [Accessed March 24, 2017]
- D'ANDREA, A., & FERNIE, K. (2013). *3D Digitisation of Icons of European Architectural and Archaeological Heritage* (Final No. D6.1: Report on Metadata and Thesaurii) (p. 45). European Commission's ICT Policy Support Programme. Retrieved from <http://www.3Dicons-project.eu/eng/Resources/D6.1-Report-on-Metadata-and-Thesaurii>
- HUMANITIES HERITAGE 3D VISUALIZATION: THEORY AND PRACTICE. (n.d.). Available from: <https://www.neh.gov/divisions/odh/institutes/humanities-heritage-3d-visualization-theory-and-practice>. [Accessed March 24, 2017]
- JOHNSTON, L. R., ASSOCIATION OF COLLEGE AND RESEARCH LIBRARIES, & AMERICAN LIBRARY ASSOCIATION. (2017). *Curating reserach data. a handbook of current practice with 30 case studies contributed by practitioners in the field Volume two Volume two*. Accessed Available from: http://www.ala.org/acrl/sites/ala.org.acrl/files/content/publications/booksanddigitalresources/digital/9780838988633_crd_v2_OA.pdf
- MARTORELLI, M., PENZA, C., & SPERANZA, D. (2014). Digital Photogrammetry for Documentation of Maritime Heritage. *Journal of Maritime Archaeology*, 9(1), 81–93. Available from: <https://doi.org/10.1007/s11457-014-9124-x>
- MOLLOY, B., WISNIEWSKI, M., LYNAM, F., O'NEILL, B., O'SULLIVAN, A., & PEATFIELD, A. (2016). Tracing edges: A consideration of the applications of 3D modelling for metalwork wear analysis on Bronze Age bladed artefacts. *Journal of Archaeological Science*, 76, 79–87. Available from: <https://doi.org/10.1016/j.jas.2016.09.007>

- NEH ANNOUNCES PROTECTING OUR CULTURAL HERITAGE. (2016, March 7). Available from: <https://www.neh.gov/news/press-release/2016-03-09>. [Accessed March 24, 2017]
- NIVEN, L., STEELE, T. E., FINKE, H., GERNAT, T., & HUBLIN, J.-J. (2009). Virtual skeletons: using a structured light scanner to create a 3D faunal comparative collection. *Journal of Archaeological Science*, 36(9), 2018–2023. Available from: <https://doi.org/10.1016/j.jas.2009.05.021>
- PROJECT HISTORY | MAYAARCH3D. (n.d.). Available from: <http://www.mayaarch3d.org/language/en/about/project-history/> [Accessed March 24, 2017]
- SOUSA ARCHIVES MUSIC INSTRUMENT DIGITAL IMAGE AND 3D MODEL COLLECTION. (n.d.). Available from: <http://imagesearchnew.library.illinois.edu/cdm/landingpage/collection/sousa>. [Accessed March 24, 2017]
- STAFF, I. B. R. (2015, September 14). ISU professor earns national grant to preserve artifacts with 3D scans. Available from: <http://idahobusinessreview.com/2015/09/14/isu-professor-earns-national-grant-to-preserve-artifacts-with-3d-scans/>. [Accessed March 24, 2017]
- WACHOWIAK, M. J., & KARAS, B. V. (2013). 3D Scanning and Replication for Museum and Cultural Heritage Applications. *Journal of the American Institute for Conservation Journal of the American Institute for Conservation*, 48(2), 141–158.
- WHAT IS DIGITAL CURATION? | DIGITAL CURATION CENTRE. (n.d.). Accessed March 24, 2017, Available from: <http://www.dcc.ac.uk/resources/briefing-papers/introduction-curation/what-digital-curation>. [Accessed March 24, 2017]
- YASTIKLI, N. (2007). Documentation of cultural heritage using digital photogrammetry and laser scanning. *Journal of Cultural Heritage*, 8(4), 423–427. <https://doi.org/10.1016/j.culher.2007.06.003>

¹ Jennifer Moore, Washington University in St. Louis, 1 Brookings Drive, Campus Box 1169, St. Louis, MO 63130. Mail: j.moore@wustl.edu.

² Hannah Scates Kettler, University of Iowa, 100 Main Library, Digital Research and Publishing, Iowa City, IA 52242

³ Community Standards for 3D Preservation Survey conducted in 2017 underscores the continued lack of community and standard practice. <https://mfr.osf.io/render?url=https://osf.io/tcn6h/?action=download%26mode=render>

⁴ <http://archaeologydataservice.ac.uk/>

⁵ <http://3Dicons-project.eu/>

⁶ <http://www.londoncharter.org/>

⁷ <https://www.digitalantiquity.org/>

⁸ <http://pro.carare.eu/doku.php?id=support:metadata-schema>

⁹ <http://augmentedasburypark.com/>

¹⁰ <http://humanorigins.si.edu/evidence/3d-collection/artifact>

¹¹ <http://vwahl.soic.indiana.edu/villa/index.php>

¹² <http://www.meshlab.net/>

¹³ <https://gomeasure3d.com/hdi/advance/>

¹⁴ <http://3Dhop.net/index.php>

¹⁵ Such conferences and institutes like the Digital Library Federation Forum (<https://forum2017.diglib.org/>), the NEH funded Advanced Challenges in Theory and Practice in 3D Modeling of Cultural Heritage Sites

(<https://www.neh.gov/divisions/odh/institutes/advanced-challenges-in-theory-and-practice-in-3D-modeling-cultural-heritage>).

¹⁶ CS3DP Google Group <https://groups.google.com/forum/#!forum/community-standards-for-3D-data-preservation-cs3Dp>

¹⁷ Community Standards for 3D Preservation Survey conducted in 2017 underscores the continued lack of community and standard practice.

<https://mfr.osf.io/render?url=https://osf.io/tcn6h/?action=download%26mode=render>